М. В. КОЗЛОВ

ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ В ПРИМЕРАХ И ЗАДАЧАХ

ДОПУЩЕНО ГОСУДАРСТВЕННЫМ КОМИТЕТОМ СССР
ПО НАРОДНОМУ ОБРАЗОВАНИЮ В КАЧЕСТВЕ УЧЕБНОГО
ПОСОБИЯ ДЛЯ СТУДЕНТОВ, ОБУЧАЮЩИХСЯ
ПО СПЕЦИАЛЬНОСТИ «МАТЕМАТИКА»

Издательство
Московского университета
1990

ББК 22.171
К59
УДК 519.21

Рецензенты:

кафедра теории вероятностей
и математической статистики
Ленинградского государственного университета
(заведующий кафедрой профессор В. В. Петров),
академик Ю. В. Прохоров

Козлов М. В.

К59 Элементы теории вероятностей в примерах и задачах. — М.: Изд-во МГУ, 1990 г. — 344 с.: ил.
ISBN 5—211—00312—8.

Основы теории вероятностей излагаются в форме примеров и задач, к которым в тексте приведены подробные решения. Уровень сложности колеблется в широком диапазоне: от тренировочных задач на усвоение понятий до маленьких исследований, могущих служить началом курсовой работы. Всего примеров и задач около 450. Принцип изложения — от частных моделей к общим понятиям — направлен на развитие у читателя вкуса и навыков к самостоятельному научному творчеству. Для освоения материала достаточно владения началами математического анализа.

К $\frac{1602090000(4309000000)—053}{077(02)—90}$ 83—89 ББК 22.171

ISBN 5—211—00312—8

Оглавление

Предисловие

Начинающие изучать теорию вероятностей стоят перед двумя проблемами. Во-первых, в понимании предмета исключительно велика роль интуитивной базы. Теорию вероятностей можно сравнить в этом отношении с теоретической механикой. Однако, если знакомство с элементами механики начинается со школьной скамьи и подкрепляется собственным практическим опытом, то, к сожалению, понятие вероятности, которое с успехом могло бы быть усвоено в средней школе, появляется лишь в вузовских курсах. В результате слушатели вынуждены следовать за формально-математическими построениями теории с еще не сложившимся интуитивным базисом. Неудивительно, что общий уровень вероятностно-статистической культуры остается фантастически низким. Процесс «дозревания» даже для хороших студентов затягивается в сравнении с другими математическими дисциплинами. По-видимому, настала пора обеспечить пакетами программ курсы теории вероятностей, чтобы можно было на собственном опыте работы с моделированными данными получить представление о характере проявления вероятностных закономерностей. Пока приходится опираться на традиционный подход, связанный с обращением к азартным играм, урновым схемам и т. п.

Вторая проблема изучающего теорию вероятностей связана с используемым математическим аппаратом. Дело в том, что уже самые простые модели приводят к вопросам, не вмещающимся в рамки элементарной теории. Так, формулируемая на уровне комбинаторной вероятности задача о разорении игрока для своего логически точного обоснования требует обращения к несчетным пространствам элементарных событий и построению вероятностной меры на них. Следствием указанных проблем часто оказывается то, что теория вероятностей превращается в полуфизическую науку, причем студент, как правило, не приобретает достаточного умения решать типовые задачи ни на уровне содержательных рассуждений, ни тем более на должном уровне строгости.

Основной принцип построения предлагаемого пособия — от частных моделей к общим понятиям. Представляется важным, чтобы появление новых понятий было по возможности мотивировано конкретными запросами решаемых задач и диктовалось внутренней логикой развивающейся теории. В этой связи неоценимую помощь оказывает анализ простых вероятностных моделей, служащих прототипами более общих. Поэтому уже в первом параграфе появляется модель простого симметричного блуждания, во втором — задача о разорении игрока, в третьем — пуассоновский процесс и броуновское движение и т. д. (Я стремился к тому, чтобы вводимые понятия и методы сразу «работали». Здесь неизбежны определенные издержки. Так, математическое ожидание появляется только во второй главе книги и тем самым, если судить по названиям глав, не входит в «начальные понятия теории».) Условная вероятность, излагаемая в § 5, увязывается с началами теории марковских цепей и применением к статистике.

Один из лейтмотивов первой главы находит свое явное воплощение в § 6, где речь идет об аксиоматике Колмогорова и продолжении вероятностной меры. Тот факт, что процесс определения вероятности событий с формальной точки зрения подобен определению понятия площади фигур, как мне представляется, должен быть освещен уже на начальном периоде обучения, особенно если речь идет о математической аудитории. Уровень строгости здесь может быть разный. В § 6, отмеченном как более трудный звездочкой, можно при первом чтении ознакомиться с несколькими начальными страницами, где разбираются свойства аксиоматического объекта теории — вероятностного пространства. По мере продвижения можно обращаться к § 6 за справками по вопросам построения вероятностной меры в различных пространствах элементарных событий.

Не будет преувеличением сказать, что в идейном отношении книга примыкает к классическому направлению теории вероятностей — предельным теоремам. Однако, несмотря на довольно значительный объем пособия, здесь не делается попытка сколько-нибудь полного изложения традиционных вопросов этого раздела теории. Задача книги — пробудить интерес к проблеме, дать первое представление о методах исследования, сохраняя ситуацию возможно более простой. С этой точки зрения центральное место занимает третья (и последняя) глава книги. Она начинается со схемы Бернулли и простого случайного блуждания (§ 11, 12). Несмотря на элементарный характер модели, она позволяет затронуть такие важные и интересные вопросы, как теорему о сходимости к устойчивым законам, сходимость к процессу броуновского движения, броуновскому мосту, броуновской экскурсии. Приводится интересная связь с ветвящимися процессами. Изучение предельного поведения биномиального распределения в § 12 сопровождается родственными задачами на сходимость к нормальному закону, служащими подготовкой к § 13, где рассматриваются вопросы сходимости вероятностных распределений, преобразований Лапласа и характеристических функций. Итог — центральная предельная теорема для независимых слагаемых. В § 14 рассматривается ряд моделей марковских цепей: неоднородное простое блуждание, ветвящиеся процессы, процессы с иммиграцией, ветвящийся процесс в случайной среде. На этих примерах демонстрируются постановки задач, обычно рассматриваемых в теории марковских цепей. Читатель подготавливается к восприятию предельной теоремы о переходных вероятностях и дискретной теоремы восстановления, чем заканчивается книга. В продолжение общей линии здесь также устанавливаются условные предельные теоремы для ветвящихся процессов, включая сходимость конечномерных распределений.

В главе второй излагается более традиционный материал о математическом ожидании, дисперсии, ковариациях, условном математическом ожидании. Указано применение к построению оптимальной в среднеквадратичном оценки и к наилучшему несмещенному оцениванию. Развернутое изложение теории интеграла и математического ожидания содержит § 10, отмеченный звездочкой.

Систематическое изучение материала § 10 будет полезно для математически более подготовленного читателя. Для тех, кого не беспокоят вопросы математической строгости, § 10 и ранее упомянутый § 6 могут служить для справок.

Содержание книги составляют свыше 450 примеров и задач, к последним в тексте приведены достаточно подробные решения. Примеры и задачи связаны более-менее единой «сюжетной линией» и служат в первую очередь не отработке типовых приемов решения задач, а направлены на развитие у читателя вкуса и навыков к активному, творческому освоению предмета. Для начинающего можно предложить следующий метод работы: примеры внимательно прочитываются с карандашом и бумагой, решение задачи разбирается лишь после неудачных попыток решить ее самостоятельно. Следует иметь в виду, что во многих случаях задачи расположены по гнездовому типу, а весь материал довольно тесно завязан. Возможно, однако, и выборочное решение задач, что облегчается ссылками на предшествующий материал, необходимый для решения данной задачи. Уровень задач весьма неоднороден, наиболее трудные задачи отмечены звездочкой.

Книга основана на традициях преподавания теории вероятностей на механико-математическом факультете Московского университета, автор использовал научные и методические достижения, содержащиеся в различных изданиях, свой собственный опыт преподавания и руководства курсовыми и дипломными работами. Список литературы в конце книги не претендует на полноту, ссылки на научную периодику не приведены ввиду того, что заимствования относятся к элементарным частным случаям.

Выражаю признательность А. Д. Вентцелю, прочитавшему книгу в рукописи и сделавшему ряд ценных замечаний. Автор благодарен кафедре теории вероятностей и математической статистики Ленинградского университета и ее заведующему В. В. Петрову за внимательное отношение к рукописи. Мне приятно выразить благодарность Ю. В. Прохорову за поддержку.

Я благодарен главному редактору издательства А. И. Камзолову и заведующему редакцией С. И. Зеленскому за творческое обсуждение плана книги. Оформлением рукописи я обязан Т. В. Нистратовой, а также студентам и аспирантам кафедры математической статистики и случайных процессов, всем им приношу искреннюю благодарность.

Мне посчастливилось со студенческих времен испытать творческое влияние дорогого учителя Андрея Николаевича Колмогорова, и эта книга — скромная дань его памяти.

Глава I

НАЧАЛЬНЫЕ ПОНЯТИЯ

§ 1. ВЕРОЯТНОСТЬ В КЛАССИЧЕСКОЙ СХЕМЕ

Что такое случайность, как она проявляется, в чем ее природа? Теория вероятностей предлагает математическую модель, оправдавшую себя при исследовании многих явлений и процессов, не поддающихся чисто детерминированным методам анализа. Наглядное представление о случайных событиях можно получить, рассматривая различные опыты с подбрасыванием монет, игральных костей, раздачей игральных карт и др. Полная неопределенность в исходе такого рода опыта, казалось бы, отвергает возможность установления здесь каких-либо закономерностей. Однако это не так. Хорошо известно, что при многократном повторении опытов наблюдается так называемая устойчивость частот появления исходов. Ниже приведены результаты эксперимента, имитирующего подбрасывание «идеальной» монеты. Всего проведено 10 000 испытаний и вычислены отношения числа появлений герба к числу испытаний после 100, 200, ..., 900, 1000, 2000, ..., 10 000 испытаний:

0,410 0,400 0,450 0,467 0,476 0,495 0,492 0,495 0,485 0,484
0,489 0,491 0,493 0,492 0,507 0,491 0,497 0,499 0,498.

Эти и многие другие экспериментальные данные приводят к заключению, что при неограниченном продолжении испытаний подобного рода частота исхода стремится к 1/2. Подбрасывая «идеальную» игральную кость, можно обнаружить, что частоты каждого из 6 исходов приближаются с ростом числа испытаний к 1/6. В общем случае если проводить многократно опыт с N исходами, ни один из которых не имеет какого-либо преимущества в реализации перед другими, то частоты его исходов будут приближаться к $1/N$. Таков *эмпирический закон устойчивости частот*. Как и любые другие эмпирические закономерности, он представляет собой некоторую идеализацию действительных отношений в природе. Практически невозможно обеспечить полную симметрию исходов опыта, так же как и нельзя провести неограниченную серию испытаний. Преимущество математического анализа явлений заключается в том, что реальным закономерностям отвечают логически точные утверждения относительно математической модели явления. Именно так обстоит дело и с теорией вероятностей, один из замечательных результатов которой — *закон больших чисел* — утверждает, что в модели бесконечной последовательности испытаний частоты исходов сходятся к некоторым предельным значениям.

Исходы опыта в теории вероятности называют *элементарными событиями*. Результат опыта — какое-нибудь одно элементарное событие. Как правило, интерес представляет не отдельное элементарное событие, а какая-либо его характеристика. Например, при извлечении из колоды игральной карты можно интересоваться ее мастью или значением. Говорим, что произошло событие {извлечена трефа}, если была получена одна из карт: туз треф, король треф, ... Событие {извлечен туз} означает, что получена одна из карт: туз пик, туз треф, ... Воообще под *событием* понимают логическое объединение некоторой совокупности элементарных событий. Считается, что данное событие происходит, если происходит любое из составляющих его элементарных событий, и не происходит — в противном случае. К событиям можно применять различные логические операции: $\vee$ (или), & (и), $\neg$ (не), $\to$ (вытекает). Событие $\mathcal{A} \vee \mathcal{B}$ происходит, когда происходит хотя бы одно из событий $\mathcal{A}$ или $\mathcal{B}$, событие $\mathcal{A} \& \mathcal{B}$ — когда происходят оба события $\mathcal{A}$ и $\mathcal{B}$, событие $\neg\mathcal{A}$ — если не происходит $\mathcal{A}$. Говорят, что $\mathcal{A} \to \mathcal{B}$, если событие $\mathcal{B}$ происходит всегда, когда происходит $\mathcal{A}$. Отметим, что знак логического умножения & обычно опускают и пишут $\mathcal{A}\mathcal{B}$ вместо $\mathcal{A} \& \mathcal{B}$.

1.1. Задача. Пусть $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ — произвольные события. Записать выражения для событий: (I) из $\mathcal{A}$, $\mathcal{B}$ и $\mathcal{C}$ произошло только $\mathcal{A}$; (II) произошли $\mathcal{A}$ и $\mathcal{B}$, но $\mathcal{C}$ не произошло; (III) все три события произошли; (IV) произошло хотя бы одно из $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$; (V) произошло только одно из событий $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$; (VI) произошло хотя бы два события; (VII) произошло только два события из $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$; (VIII) ни одно из событий $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ не произошло; (IX) произошло не более двух событий.

[I) $\mathcal{A}(\neg\mathcal{B})(\neg\mathcal{C})$, (II) $\mathcal{A}\mathcal{B}(\neg\mathcal{C})$, (III) $\mathcal{A}\mathcal{B}\mathcal{C}$, (IV) $\mathcal{A} \vee \mathcal{B} \vee \mathcal{C}$, (V) $(\mathcal{A}(\neg\mathcal{B})(\neg\mathcal{C})) \vee ((\neg\mathcal{A})\mathcal{B}(\neg\mathcal{C})) \vee ((\neg\mathcal{A})(\neg\mathcal{B})\mathcal{C})$, (VI) $(\mathcal{A}\mathcal{B}) \vee (\mathcal{A}\mathcal{C}) \vee (\mathcal{B}\mathcal{C})$, (VII) $(\mathcal{A}\mathcal{B}(\neg\mathcal{C})) \vee (\mathcal{A}(\neg\mathcal{B})\mathcal{C}) \vee ((\neg\mathcal{A})\mathcal{B}\mathcal{C})$, (VIII) $(\neg\mathcal{A})(\neg\mathcal{B})(\neg\mathcal{C})$, (IX) $(\neg\mathcal{A}\mathcal{B}\mathcal{C})$.]

Классическое исчисление вероятностей (XVII век) рассматривало такие опыты, исходы которых обладают определенной симметрией, позволяющей рассчитывать на одинаковую возможность появления любого из исходов. Практически это означает, что при повторении опыта частоты всех его N исходов приближаются с ростом числа испытаний к N^{-1}. Для события $\mathcal{A}$, состоящего из M элементарных событий, отсюда следует, что частота его появления (равная сумме частот составляющих его элементарных событий) стремится с ростом числа испытаний к M/N. Отношение числа M элементарных событий, составляющих данное событие $\mathcal{A}$, к полному числу N элементарных событий опыта с *равновозможными исходами* называется *вероятностью* события $\mathscr{A}$.

1.2. Задача. Три письма раскладываются случайно по трем конвертам с адресами. Найти вероятность того, что хотя бы одно письмо попадет в свой конверт.

[Представим всевозможные разложения графом, имеющим вид трехъярусного дерева (рис. 1). Три занумерованных ребра, выходящих из его корня, соответствуют возможным номерам писем, помещенных в конверт с номером 1. Из конца каждого ребра первого яруса выходят по два ребра с возможными номерами писем, помещенных в конверт номер 2. Из концов ребер второго яруса выходит по одному ребру с номером письма, помещенного в конверт номер 3. Из шести ветвей дерева четыре отвечают интересующему нас событию, так что искомая вероятность равна 4/6].

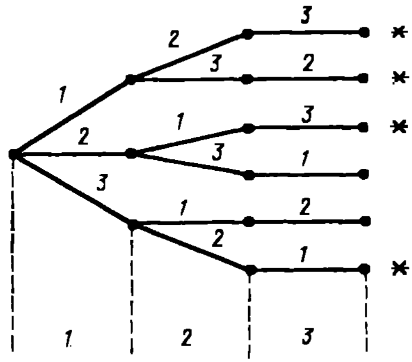

Рис. 1

1.3. Задача. Два шара размещаются случайно по трем ящикам без ограничений на число шаров в ящике. Найти вероятность того, что шары попадут в разные ящики.

[Занумеруем шары и ящики и опишем результат разложения парой (i_1, i_2), i_k — номер ящика, куда помещен шар с номером k. Всего имеется $3^2=9$ элементарных событий $\omega=(i_1, i_2)$, из которых (1,1), (2,2), (3,3) отвечают попаданию шаров в один ящик, а оставшиеся 6 элементарных событий составляют нужное событие, его вероятность равна 6/9.]

1.4. Задача. Показать, что из чисел $1, 2, \ldots, N$ можно составить N^n различных последовательностей $\omega=(i_1, i_2, \ldots, i_n)$, из которых $N(N-1)\ldots(N-n+1)$ последовательностей состоят из различных чисел: $i_k \neq i_l$ при $k \neq l$ $(n \leqslant N)$.

[Обозначим $\alpha_{n,N}$ число всех последовательностей ω, $\beta_{n,N}$ — число последовательностей с попарно различными членами. Разбивая те и другие последовательности на подмножества, состоящие из последовательностей вида $(1, i_2, \ldots, i_n)$, $(2, i_2, \ldots, i_n)$, ..., $(N, i_2, \ldots, i_n)$, получаем рекуррентные соотношения

$$\alpha_{n,N}=N\alpha_{n-1,N},\quad \beta_{n,N}=N\beta_{n-1,N-1},$$

из которых с учетом $\alpha_{1,N}=N$, $\beta_{1,k}=k$ находим ответ.]

1.5. Задача. На экзамене n студентов по очереди вытаскивают билеты, из которых M — «легкие», а остальные $N-M$ — «трудные». Какова вероятность, что студенту, пришедшему на экзамен k-м, достанется легкий билет, если: (I) вытащенный билет назад не возвращается; (II) возвращается.

[Пусть для определенности легкие билеты имеют номера от 1 до M. В случае (I) представим элементарные события опыта строчками $\omega=(i_1, i_2, \ldots, i_n)$, где i_k — номер билета, полученного k-м студетом, $k=1, 2, \ldots, n$. Всего имеется $N(N-1)\ldots(N-n+1)$ различных последовательностей ω (см. 1.4). Последовательностей $\omega=(i_1, \ldots, i_{k-1}, j, i_{k+1}, \ldots, i_n)$ с фиксированным j имеется $(N-1)(N-2)\ldots(N-1-(n-1)+1)$, а с любым $j\leqslant M$ число последовательностей в M раз больше. Отсюда получаем искомую вероятность $M(N-1)\ldots(N-n+1)(N(N-1)\ldots(N-n+1))^{-1}=M/N$. В случае (II) имеется N^n элементарных событий, из которых MN^{n-1} составляют интересующее событие, что приводит к той же вероятности M/N, что и для (I).].

1.6. Задача. У человека имеется n ключей, из которых только один подходит к замку. Он испытывает их в случайном порядке, пока замок не откроется. Какова вероятность, что это произойдет при k-м испытании, если: (I) опробованный и не подошедший ключ откладывается; (II) возвращается к остальным ключам.

[(I)Будем испытывать ключи и после того, как замок откроется (что не меняет ответа). Занумеровав ключи, опишем порядок извлечения последовательностью $(i_1, i_2, \ldots, i_n)$ из n различных чисел от 1 до n Всего имеется $n(n-1)\ldots 2\cdot 1=n!$ таких последовательностей (см. 1.4 при $N=n$). Пусть для определенности нужный ключ имеет номер 1. Последовательности вида $(i_1, \ldots, i_{k-1}, 1, i_{k+1}, \ldots, i_n)$ отвечают событию, что замок открыт при k-м испытании, а число таких последовательностей равно $(n-1)!$. Искомая вероятность равна $(n-1)!/n!=1/n$ при любом k от 1 до n. В случае (II) число испытаний теоретически неограничено, поэтому мы поставим какую-либо границу N для числа испытаний и найдем вероятность того, что замок откроется до N-го испытания. Чтобы можно было воспользоваться схемой опыта с равновозможными исходами, представим себе, что испытания продолжаются и после того, как замок открылся, и, следовательно, результат опыта описывается последовательностью $(i_1, i_2, \ldots, i_N)$, где каждое i_k принимает любое значение от 1 до n, $k=1, 2, \ldots, N$. Всего имеется n^N последовательностей, из них событию {замок впервые открылся при k-м испытании} отвечают последовательности вида $(i_1, \ldots, i_{k-1}, 1, i_{k+1}, \ldots, i_N)$, где $i_1\neq 1, \ldots, i_{k-1}\neq 1$, $i_{k+1}, \ldots, i_N$ — произвольные. Отсюда находим ответ: $(n-1)^{k-1}n^{N-k}n^{-N}=(1-1/n)^{k-1}\cdot 1/n$, который не зависит от ограничения N на число испытаний, если только для данного k выбрать $N\geqslant k$. Кажется вполне логичным принять $p_k=(1-1/n)^{k-1}\cdot 1/n$, $k=1, 2, \ldots$, за ответ в задаче (II). Отметим, что сумма ряда $p_1+p_2+\ldots=1$.]

1.7. **Задача.** Пусть n человек выстраиваются случайным образом в очередь. Какова вероятность, что между X и Y будут стоять ровно r человек.

[Присвоив каждому из n людей номер, представим любое расположение в очередь последовательностью $(i_1, i_2, \ldots, i_n)$, где i_k —

номер человека, занимающего в очереди k-е место, $k=1, 2, \ldots, n$. Всего имеется $n(n-1)\ldots 2\cdot 1=n!$ различных последовательностей $(i_1, i_2, \ldots, i_n)$ (см. 1.4 при $N=n$). Обозначим m_X, m_Y номера X и Y. Последовательностей вида $(i_1, \ldots, i_{k-1}, m_X, i_{k+1}, \ldots, i_{k+r}, m_Y, i_{k+r+2}, \ldots, i_n)$ при фиксированных k и r имеется $(n-2)!$. Перебирая всевозможные $k=1, 2, \ldots, n-r-1$ и меняя местами X и Y, получим $2(n-r-1)(n-2)!$ различных расстановок, удовлетворяющих условию. Искомая вероятность равна

$$2(n-r-1)(n-2)!/n!=2(1-r/(n-1))n^{-1},\ r=0, 1, \ldots, n-2.]$$

Пусть $\mathscr{X}=\{x_1, x_2, \ldots, x_N\}$ — произвольное множество из N элементов. Последовательность $\omega=(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$, составленную из n различных элементов множества $\mathscr{X}$, называют *выборкой объема n без возвращения* из множества $\mathscr{X}$, $n\leqslant N$. Множество $\Omega=\{\omega\}$ всех выборок без возвращения объема n в соответствии с 1.4 состоит из $N(N-1)\ldots(N-n+1)$ элементов. При $n=N$ выборку без возвращения объема N называют *перестановкой* или *упорядочиванием* элементов множества $\mathscr{X}$. Всего имеется $N(N-1)\ldots 2\cdot 1=N!$ различных перестановок элементов множества $\mathscr{X}$. *Случайным выбором без возвращения* объема n из множества $\mathscr{X}$ называется опыт с $N(N-1)\ldots(N-n+1)$ равновозможными исходами $\omega=(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$. *Выборкой объема n с возвращением* из множества $\mathscr{X}$ называют произвольную последовательность $\omega'=(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$, где x_{i_k}, $k=1, 2, \ldots, n$, — возможно повторяющиеся элементы множества $\mathscr{X}$. В соответствии с 1.4 в множестве $\Omega'=\{\omega'\}$ имеется N^n элементов, и число n может быть взято любым. *Случайным выбором с возвращением* объема n из множества $\mathscr{X}$ называют опыт с N^n равновозможными исходами, представленными элементами множества Ω'.

1.8. Задача. Показать, что N-элементное множество имеет $C_N^n=N(N-1)\ldots(N-n+1)/n!$ n-элементных подмножеств.

[Произвольную выборку без возвращения объема n из множества $\mathscr{X}=\{x_1, x_2, \ldots, x_N\}$ можно образовать, взяв подмножество $\mathscr{X}'\subseteq\mathscr{X}$ из n элементов и упорядочив его. Так как подмножество из n элементов можно упорядочить $n!$ различными способами, то, обозначив C_N^n число всех n-элементных подмножеств $\mathscr{X}'$, получаем равенство $C_N^n\cdot n!=N(N-1)\ldots(N-n+1)$, откуда находим C_N^n.]

В комбинаторике подмножество конечного множества называют *сочетанием* из его элементов. Число

$$C_N^n=\frac{N(N-1)\ldots(N-n+1)}{n!}=\frac{N!}{n!\,(N-n)!}$$

называют числом сочетаний из N элементов по n элементов. Положим дополнительно $C_N^0=1$, $0!=1$, $C_N^k=0$ при $k>N$ и $k<0$. Отметим очевидное тождество $C_N^n=C_N^{N-n}$.

1.9. Задача. Установить тождество $C_N^n = C_{N-1}^{n-1} + C_{N-1}^n$, разбивая все сочетания из N элементов по n на содержащие и не содержащие некоторый фиксированный элемент.

[Подмножеств $\{x_{i_1}, \ldots, x_{i_n}\} \subseteq \{x_1, x_2, \ldots, x_N\}$, не содержащих элемент x_1, имеется столько же, сколько существует n-элементных подмножеств множества $\{x_2, \ldots, x_N\}$, т. е. C_{N-1}^n. Всякое подмножество $\{x_{i_1}, \ldots, x_{i_n}\}$, содержащее элемент x_1, можно образовать, присоединяя этот элемент к любому $(n-1)$-элементному подмножеству множества $\{x_2, \ldots, x_N\}$, так что их число равно C_{N-1}^{n-1}. Складывая C_{N-1}^{n-1} и C_{N-1}^n, получаем число всех n-элементных подмножеств множества $\{x_1, x_2, \ldots, x_N\}$, т. е. C_N^n.]

1.10. Задача. Показать, что имеется C_N^n строчек $(\delta_1, \delta_2, \ldots, \delta_N)$ из 0 и 1, содержащих ровно n единиц. Вывести отсюда тождество

$$C_N^0 + C_N^1 + \ldots + C_N^N = 2^N.$$

[Поставим в соответствие каждому подмножеству множества $\{x_1, x_2, \ldots, x_N\}$, включая пустое, строчку $(\delta_1, \delta_2, \ldots, \delta_N)$, где $\delta_i = 1$, если x_i входит в подмножество, и $\delta_i = 0$ — в противном случае. Это соответствие взаимно однозначно. Остается воспользоваться задачей 1.8.]

1.11. Пример (симметричное случайное блуждание). Рассмотрим модель перемещения некоторой частицы по целым точкам прямой, предполагая, что каждую единицу времени частица сдвигается на единичное расстояние вправо или влево и направление движения определяется случайным выбором с возвращением из множества $\{+1, -1\}$. Каждой выборке $\omega = (x_1, x_2, \ldots, x_n)$, $x_i = \pm 1$, $i = 1, 2, \ldots, n$, поставим в соответствие траекторию движения $(s_0, s_1, \ldots, s_n)$, где $s_0 = a$ — начальное положение частицы, $s_k = a + x_1 + \ldots + x_k$ — ее положение в момент времени k, $k = 1, 2, \ldots, n$. Описанную модель называют *простым случайным блужданием*, это блуждание *симметрично* в том смысле, что выбор направления шага каждый раз происходит с равными вероятностями 1/2 (см. 1.5 (II) при $N=2$, $M=1$).

Между элементарными событиями ω и траекториями $(s_0, s_1, \ldots, s_n)$ при фиксированном начальном положении s_0 имеется взаимно однозначное соответствие. Это позволяет при вычислении вероятностей различных событий, описывающих поведение траектории блуждания, производить подсчет траекторий вместо элементарных событий. При изучении случайного блуждания очень полезными оказываются геометрические соображения, основанные на представлении траектории блуждания ломаной на координатной плоскости с вершинами (k, s_k), $k = 0, 1, 2, \ldots, n$ (рис. 2). Назовем эти ломаные и любые их отрезки путями. Назовем длиной пути длину соответствующего ему временного интервала. Введем обозначения: $N_{n,k}$ для числа путей, соединяющих начало координат (0, 0) с точкой (n, k), $p_{n,k} = N_{n,k} 2^{-n}$.

В задачах 1.12—1.19 изучаются различные свойства траекторий симметричного случайного блуждания.

1.12. Задача. Показать, что $N_{n,k}=C_n^{(n+k)/2}$, если k и n одинаковой четности, $|k|\leqslant n$ и $N_{n,k}=0$ в противном случае.

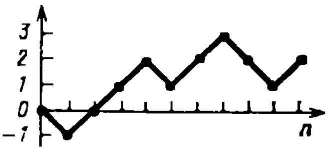

Рис. 2

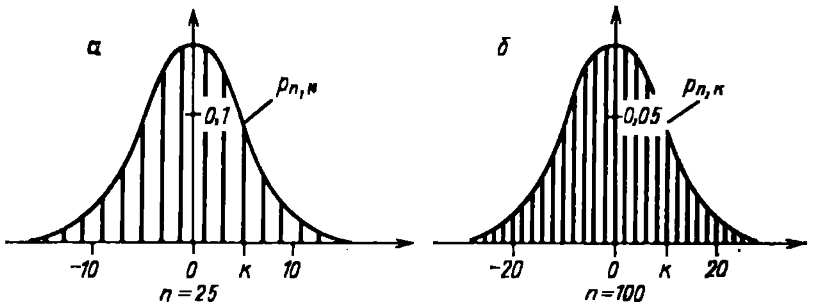

Рис. 3

[Пусть $s_n=x_1+\ldots+x_n=k$ и m — число единиц в последовательности $(x_1, \ldots, x_n)$. Тогда m и k связаны уравнением $m\cdot(+1)+(n-m)\cdot(-1)=k$, откуда $m=(n+k)/2$, если n и k одной четности. Ввиду 1.10 имеется C_n^m последовательностей $(x_1, \ldots, x_n)$, содержащих m единиц, так что $N_{n,k}=C_n^m$, $m=(n+k)/2$.]

На рис. 3 вероятности $p_{n,k}=N_{n,k}2^{-n}$ представлены вертикальными отрезками длины $p_{n,k}$ с основаниями в точках k оси абсцисс. Как видно, подавляющее большинство траекторий оканчивается в точках k, небольших сравнительно с n. Обращает на себя внимание то, что огибающие, проведенные через верхние концы отрезков для случаев $n=25$ и $n=100$, практически не отличаются друг от друга (при выбранном соотношении двух масштабов графиков).

1.13. Задача. Доказать следующий *принцип отражения*: число путей, ведущих из точки $(0, a)$, $a>0$, в точку (n, b), $b>0$, которые достигают оси абсцисс, равно числу всех путей, ведущих из точки $(0, -a)$, полученной симметричным отражением точки $(0, a)$ относительно оси абсцисс, в точку (n, b) (рис. 4). Таким образом, имеется $N_{n,a+b}$ путей из $(0, a)$ в (n, b), достигающих оси абсцисс, и $N_{n,b-a}-N_{n,a+b}$ путей, ее не достигающих.

[Отразим отрезок пути от его начала до первой точки попадания на ось абсцисс симметрично относительно оси абсцисс (рис. 4) и поставим в соответствие каждой траектории $(s_0, s_1, \ldots, s_n)$, $s_0=a$, $s_1>0, \ldots, s_{l-1}>0$, $s_l=0$, траекторию $(-s_0, -s_1, \ldots, -s_{l-1},$

$s_t, \ldots, s_n$). Это взаимно однозначное соответствие между путями рассматриваемого класса и всеми путями из $(0, -a)$ в (n, b), что и доказывает принцип отражения.]

1.14. Задача. Показать, что число положительных путей длины $2n$, ведущих из $(0, 0)$ в $(2n, 2k)$: $s_0=0, s_1>0, s_2>0, \ldots$ $\ldots, s_{2n-1}>0, s_{2n}=2k>0$, равно $N_{2n-1,2k-1}-N_{2n-1,2k+1}$, где при $k=n$ следует положить $N_{2n-1,2n+1}=0$. Вывести отсюда, что число всех положительных путей длины $2n$: $s_0=0, s_1>0, s_2>0, \ldots, s_{2n}>0$ равно $N_{2n-1,1}=C^n_{2n-1}$, а число положительных путей длины $2n+1$ равно $2N_{2n-1,1}$.

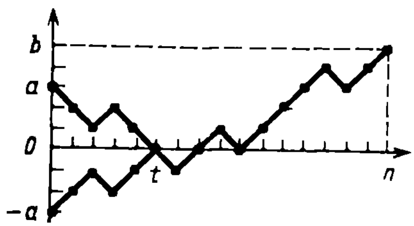

Рис. 4

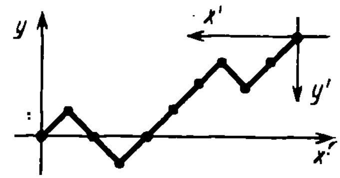

Рис. 5

[Любой положительный путь, начинающийся из точки (0, 0), проходит через точку (1, 1). Число путей длиной $2n-1$ из (1, 1) в $(2n, 2k)$, не достигающих оси абсцисс, равно ввиду 1.13 $N_{2n-1,2k-1}-N_{2n-1,2k+1}$. Суммируя по k, найдем число всех положительных путей

$$\sum_{k=1}^{n}(N_{2n-1,2k-1}-N_{2n-1,2k+1})=\sum_{k=1}^{n}N_{2n-1,2k-1}-\sum_{k=2}^{n}N_{2n-1,2k-1}=N_{2n-1,1}.$$

Любой положительный путь длиной $2n$ имеет ординату конечной точки $s_n\geqslant 2$, так как равенство $s_{2n}=1$ невозможно. Продолжая этот путь еще на один шаг, получим положительный путь длиной $2n+1$, так что число этих путей равно $2N_{2n-1,1}$.]

1.15. Задача. Показать, что имеется $2N_{2n-1,1}$ неотрицательных путей длиной $2n$ и $N_{2n-1,1}$ — длиной $2n-1$ $(s_0=0)$.

[Перенесем начало любого пути рассматриваемого вида в точку (1, 1) и, дополнив этот путь звеном, соединяющим точки (0, 0) и (1, 1), получим положительный путь на единицу большей длины. Остается воспользоваться задачей 1.14.]

1.16. Задача. Показать, что путей длиной n, у которых крайняя правая точка наивысшая: $s_n>s_k$, $k=0, 1, \ldots, n-1$, столько же, сколько имеется положительных путей, начинающихся из точки (0, 0).

[Каждый путь длиной n с единственным максимумом в концевой точке (n, s_n) рассмотрим в системе координат x', y' с началом в точке (n, s_n) и осями x', y', противоположно направленными к x, y (рис. 5) Легко видеть, что в системе (x', y') такой путь будет положительным, а соответствие между путями взаимно однозначно. Более формально: каждому элементарному событию $\omega = (x_1, x_2, \ldots, x_n)$ поставим в соответствие обращенную последовательность $\omega^* = (x_n, x_{n-1}, \ldots, x_1)$, а каждой траектории $(s_0, s_1, \ldots, s_n)$, $s_0 = 0$, $s_k = x_1 + \ldots + x_k$, $k = 1, \ldots, n$, — *двойственную траекторию* $(s_0^*, s_1^*, \ldots, s_n^*)$, $s_0^* = 0$, $s_k^* = x_n + x_{n-1} + \ldots + x_{n-k+1} = s_n - s_{n-k}$, $k = 1, \ldots, n$. В таком случае $s_k^* > 0$ при $k = 1, \ldots, n$ означает, что $s_n - s_{n-k} > 0$ при $k = 1, \ldots, n$, т. е. точка (n, s_n) — наивысшая.]

Проведенные в 1.12—1.16 расчеты позволяют найти вероятности ряда событий, описывающих поведение случайного блуждания заданной длительности. Пусть для определенности частица выходит из точки 0, а блуждание продолжается $2n$ единиц времени. Вероятность того, что в конечный момент времени $2n$ частица будет находиться в точке $2k$, равна

$$p_{2n,2k} = C_{2n}^{n+k} 2^{-2n}.$$

Введем отдельное обозначение для вероятности возвращения в начальное состояние

$$u_{2n} = p_{2n,0} = C_{2n}^{n} 2^{-2n}.$$

Вероятность того, что в течение всего времени блуждания частица будет находиться на положительной полуоси ($s_k > 0$, $k = 1, 2, \ldots, 2n$), равна

$$N_{2n-1,1} 2^{-2n} = \frac{(2n-1)!}{n!\,(n-1)!} 2^{-2n} = \frac{1}{2} C_{2n}^{n} 2^{-2n} = \frac{1}{2} u_{2n},$$

а что частица не попадет на отрицательную полуось — u_{2n}.

Если блуждание имеет длительность $m > 2n$, то, взяв любую из $p_{2n,2k} 2^{2n}$ траекторий длительности $2n$, отвечающих событию $s_{2n} = 2k$, продолжим ее всевозможными способами после момента $2n$. Получим $p_{2n,2k} 2^{2n} 2^{m-2n}$ траекторий длительности m, проходящих в момент $2n$ через точку $2k$. Соответствующая вероятность равна $p_{2n,2k}$ и не зависит от m. Классическая схема опыта с равновозможными исходами не позволяет рассматривать блуждание неограниченной длительности, однако проведенные рассуждения показывают, что $p_{2n,2k}$ логично принять за вероятность события $s_{2n} = 2k$ и в неограниченном по времени блуждании. Этот же вывод верен и для других событий, описывающих поведение конечного отрезка траектории.

1.17. Задача. Положим $\mu_n = \max\{s_0, s_1, \ldots, s_n\}$, и пусть $s_0 = 0$. Показать, что имеется: (I) $N_{n,2r-k}$ путей, таких, что $\mu_n \geqslant r$, $s_n = k$; (II) $N_{n,2r-k} - N_{n,2r+2-k}$ путей, таких, что $\mu_n = r$, $s_n = k$; (III) $\max\{N_{n,r}, N_{n,r+1}\}$ путей, таких что $\mu_n = r$, $(k \leqslant r)$.

[Все пути со свойством $\mu_n \geqslant r$, $s_n = k$ ведут из точки $(0, 0)$ в точку (n, k) и достигают горизонтальной прямой с ординатой r. Применяя принцип отражения относительно этой прямой, находим, что число этих путей равно $N_{n,2r-k}$. Разбивая множество путей $\{\mu_n \geqslant r,\ s_n=k\}$ на пути $\{\mu_n=r,\ s_n=k\}$ и на пути $\{\mu_n \geqslant r+1,\ s_n=k\}$, получаем, что число путей в $\{\mu_n=r,\ s_n=k\}$ равно $N_{n,2r-k}-N_{n,2r+2-k}$. Множество путей $\{\mu_n=r\}$ разбивается на непересекающиеся подмножества $\{\mu_n=r,\ s_n=k\}$ $k=r,\ r-1, \ldots$. Отсюда число путей в $\{\mu_n=r\}$ равно

$$\sum_{\{k\leqslant r\}} (N_{n,2r-k}-N_{n,2r+2-k}) = \sum_{\{k\leqslant r\}} N_{n,2r-k} -$$

$$-\sum_{\{k\leqslant r-2\}} N_{n,2r-k} = N_{n,r}+N_{n,r+1} = \max\{N_{n,r},\ N_{n,r+1}\},$$

где в последнем равенстве учтено, что одно из чисел $N_{n,r}$ или $N_{n,r+1}$ равно нулю.]

На рис. 6 вероятности $q_{n,r}=\max\{p_{n,r},\ p_{n,r+1}\}$ того, что траектория длиной n имеет максимум величины r, представлены при $n=25$ вертикальными отрезками длины $q_{n,r}$, подобно тому как это

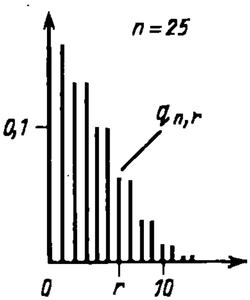

Рис. 6

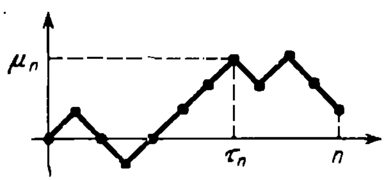

Рис. 7

сделано на рис. 3 для вероятностей $p_{n,k}$. Отличие рис. 6 от рис. 3 состоит лишь в том, что взята правая половина графика для вероятностей $p_{n,k}$, а частота отрезков удвоена за счет повторения каждого отрезка.

1.18. Задача. Обозначим τ_n момент первого достижения траекторией $(s_0, s_1, \ldots, s_n)$ своего максимума (рис. 7):

$$\tau_n = t \Leftrightarrow s_0 < s_t,\ s_1 < s_t,\ \ldots,\ s_{t-1} < s_t,\ s_t \geqslant s_{t+1},\ \ldots,\ s_t \geqslant s_n.$$

Показать, что при $0<m<n$ события $\tau_{2n}=2m$ и $\tau_{2n}=2m+1$ имеют равные вероятности $(1/2)u_{2m}u_{2n-2m}$, событие $\tau_{2n}=0$ — вероятность u_{2n}, события $\tau_{2n}=1$ и $\tau_{2n}=2n$ — равные вероятности $u_{2n}/2$.

[Имеется $(1/2)u_{2m}2^{2m}$ путей длиной $2m$, у которых концевая точка — наивысшая, и $u_{2n-2m}2^{2n-2m}$ путей длиной $2n-2m$, которые не поднимаются выше своей начальной точки (см. 1.15, 1.16). Приняв конечную точку любого пути из первой группы за начальную для любого пути из второй группы, получим путь длиной $2n$, у которого первое достижение максимума происходит в момент $2m$. Отсюда находим вероятность события $\tau_{2n}=2m$, $0<m<<n$:

$$(1/2)u_{2m}2^{2m}u_{2n-2m}2^{2n-2m}2^{-2n}=(1/2)u_{2m}u_{2n-2m}.]$$

На рис. 8 представлены вероятности $\alpha_{n,k}$ событий $\tau_n=k$ для $n=25$, 50 и значений k, взятых через один; точки $(k, \alpha_{n,k})$ соединены плавной кривой. Как видно из графика, более вероятно, что путь длиной n достигает максимума либо в своем начале, либо в конце и менее вероятно, что максимум достигается где-то посередине. Обратим также внимание на то, что огибающие верхних концов отрезков при $n=25$ и $n=50$ почти не отличаются друг от друга (при выбранном соотношении между масштабами двух графиков).

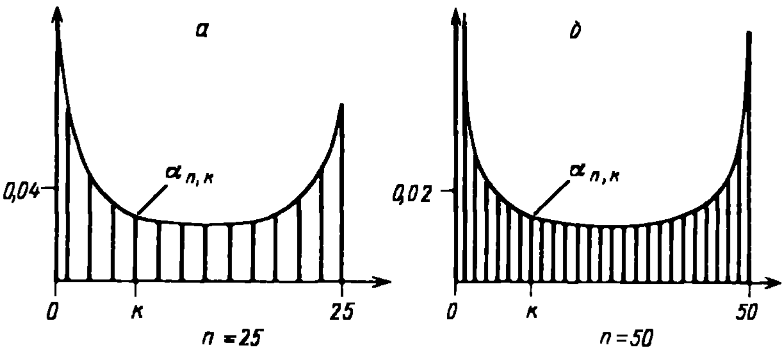

Рис. 8

1.19. Задача. Обозначим σ_n момент последнего попадания траектории $(s_0, s_1, \ldots, s_n)$ в начальное состояние s_0. Показать, что событие $\sigma_{2n}=2m$ имеет вероятность $u_{2m}u_{2n-2m}$, $m=0, 1, \ldots, n$. [Пусть $s_0=0$, возьмем любой из $u_{2m}2^{2m}$ путей длиной $2m$, для которых $s_{2m}=0$, и продолжим его любым из $(1/2)u_{2n-2m}2^{2n-2m}$ положительных путей длиной $2n-2m$, либо любым из такого же числа отрицательных путей. В результате получаем $u_{2m}2^{2m}u_{2n-2m}\cdot 2^{2n-2m}$ путей, составляющих событие $\sigma_{2n}=2m$.]

1.20. Пример (урновая модель). Представим себе урну с N занумерованными шарами и рассмотрим опыт по извлечению наугад одного шара. Если содержимое урны тщательно перемешано, то рассматриваемый опыт является, по-видимому, наиболее наглядной физической моделью, иллюстрирующей понятие равновозможности исходов. Этот опыт можно использовать как модельный и для представления опытов с неравновозможными исходами. Именно, если из N шаров M_i шаров окрашены в i-й цвет, $i=1, 2, \ldots, k$, $M_1+\ldots+M_k=N$, то появление при извлечении шара i-го цвета имеет вероятность $p_i=M_i/N$, $i=1, \ldots, k$, $p_1+\ldots+p_k=1$. Принимая цвет шара за «элементарный» исход опыта, т. е. рассматривая в качестве результата опыта не номер извлеченного шара, а его цвет, приходим к схеме опыта с конечным числом исходов, вероятности которых — некоторые рациональные числа.

Урновая модель позволяет также реализовать схему повторного случайного выбора. Предположим, что из урны один за другим извлекаются наугад шары, записываются их номера, а шары при этом в урну не возвращаются. Результат n-кратного извлечения представим последовательностью $\omega=(i_1, i_2, \ldots, i_n)$ номеров извлекаемых шаров. Легко понять, что подмножества $\{\omega : i_1=k\}$, $k=1, 2, \ldots, N$, элементарных событий этого опыта, соответствующие N различным возможностям извлечения первого шара, должны иметь одинаковые вероятности. Фиксируя k и рассматривая подмножества $\{\omega: i_1=k, i_2=l\}$, $l=1, \ldots, k-1, k+1, \ldots, N$, также приходим к заключению, что эти подмножества должны иметь одинаковые вероятности при всех l, а варьируя k, получаем, что все подмножества $\{\omega: i_1=k, i_2=l\}$, $k\neq l$, должны быть равновероятными. Продолжая этот процесс, приходим к заключению, что элементарные события $\omega=(i_1, i_2, \ldots, i_n)$ рассматриваемого опыта следует считать равновозможными. Но это и есть случайный выбор без возвращения. Случайный выбор с возвращением также реализуется в урновой модели, если каждый извлеченный шар после записи его номера возвращается назад в урну, содержимое которой тщательно перемешивается, чтобы обеспечивать при следующем извлечении равновозможность любого из N исходов.

1.21. Задача. Производится n-кратный случайный выбор с возвращением из урны с N шарами, из которых M — красные. Показать, что: (I) вероятность появления в выборке m красных шаров при испытаниях с номерами $k_1<k_2<\ldots<k_m$ равна $p^m(1-p)^{n-m}$, $p=M/N$; (II) вероятность того, что в выборке содержится m красных шаров равна

$$p_n(m; p)=C_n^m p^m (1-p)^{n-m}, \quad m=0, 1, \ldots, n, \ p=M/N.$$

[Полагая, что красные шары имеют номера от 1 до M, и замечая, что число последовательностей $(i_1, i_2, \ldots, i_n)$, таких, что $i_{k_1}\leqslant M, \ldots, i_{k_m}\leqslant M$, а остальные $i_k>M$ $(i_k\leqslant N)$, равно $M^m(N-M)^{n-m}$, что и приводит к (I). Множество элементарных событий, составляющих (II), разобъем на непересекающиеся подмножества, определяемые последовательностью $k_1<k_2<\ldots<k_m$ номеров испытаний, при которых появились красные шары. Число последовательностей $(k_1, k_2, \ldots, k_m)$ совпадает с числом m-элементных подмножеств $\{k_1, k_2, \ldots, k_m\}$ множества $\{1, 2, \ldots, n\}$ и равно C_n^m. Отсюда вероятность события (II) равна $C_n^m M^m (N-M)^{n-m} N^{-n}=p_n(m; p)$.]

На рис. 9 представлены вероятности $p_n(m, p)$ для ряда значений n и p.

1.22. Пример (размещение шаров по ящикам). Представим себе опыт, при котором каждый из n шаров помещается наудачу в любой из r ящиков. Элементарные события опыта опишем строчками $\omega=(i_1, i_2, \ldots, i_n)$, где i_k — номер ящика, куда помещен шар с номером k, $i_k=1, 2, \ldots, r$, $k=1, 2, \ldots, n$. Рассматриваемый опыт можно моделировать с помощью случайного повторного выбора: каждое размещение шара — это случайный выбор с возвращением из номеров ящиков, куда этот шар помещается. Воспользовавшись таким представлением, подсчитаем вероятность того, что заданные l ящиков будут содержать в сумме m шаров. Считая выделенные ящики «красными шарами» урновой модели, получаем из 1.21 вероятность $C_n^m p^m (1-p)^{n-m}$, $p=l/r$.

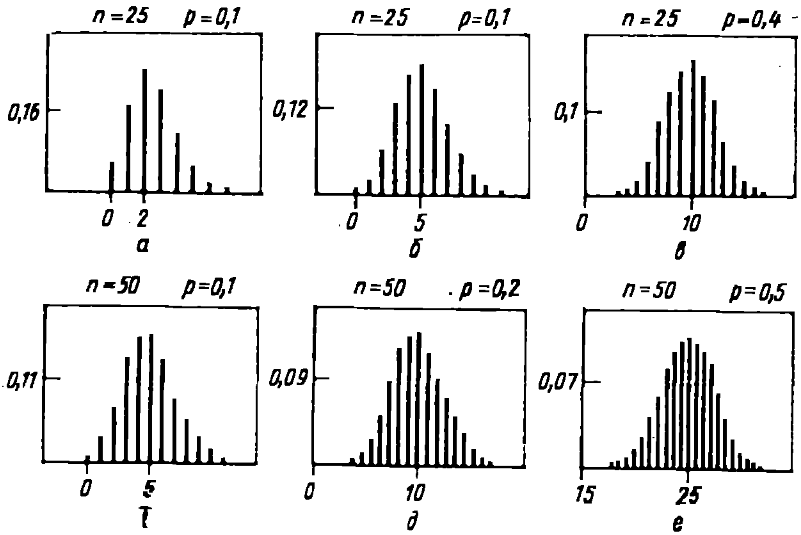

Рис. 9

1.23. Задача. Пусть n шаров случайно размещаются по n ящикам, как в 1.22 при $r=n$. Найти вероятность того, что: (I) все ящики будут заняты; (II) ровно один ящик окажется пустым ($n>1$).

[Разложения n шаров по n ящикам по одному в каждом ящике можно описать перестановкой из номеров шаров $(j_1, j_2, \ldots, j_k)$, где j_k — номер шара, помещенного в k-й ящик, $k=1, \ldots, n$. Следовательно, имеется $n!$ размещений и вероятность события (I) равна $n!/n^n$. Разложение, отвечающее событию (II), можно получить, выполняя по очереди следующие операции: выберем один из n ящиков и поместим в него любую из C_n^2 пар шаров, выберем один из $n-1$ остающихся ящиков и оставим его пустым. разложим $(n-2)!$ способами оставшиеся $n-2$ шара по одному в $n-2$ оставшихся ящика. Всего получается $nC_n^2(n-1)(n-2)!$ вариантов размещений, что приводит к вероятности $C_n^2 n!/n^n$.]

1.24. Задача. В схеме случайного выбора с возвращением из урны, содержащей M красных и $N-M$ черных шаров, найти вероятности событий: (I) первому появлению в выборке красного шара предшествует k черных; (II) r-му появлению в выборке красного шара предшествует k черных.

[Пусть красные шары имеют номера от 1 до M, а черные — от $M+1$ до N. Событие (I) состоит из таких $\omega=(i_1, i_2, \ldots, i_n)$, что $i_1>M$, $i_2>M$, $\ldots$, $i_k>M$, $i_{k+1}\leqslant M$, а остальные координаты ω — произвольные (от 1 до N), его вероятность равна $(N-M)^kM\times\times N^{n-k-1}/N^n=(1-p)^kp$, $p=M/N$. Событие (II) состоит из таких $\omega=(i_1, i_2, \ldots, i_n)$, что $i_{k+r}\leqslant M$, среди координат $i_1, i_2, \ldots, i_{k+r-1}$ ровно k имеют значения $>M$, координаты $i_{k+r+1}, \ldots, i_n$ произвольны. Число строчек $(i_1, i_2, \ldots, i_{k+r-1})$ указанного вида равно $C_{k+r-1}^k(N-M)^kM^{r-1}$, откуда искомая вероятность равна

$$C_{k+r-1}^k(N-M)^kM^{r-1}MN^{n-k-r}N^{-n}=C_{k+r-1}^kp^r(1-p)^k.]$$

1.25. Задача. В схеме размещения n шаров по r ящикам (см. 1.22) найти вероятность события {ящик с номером j содержит k_j шаров, $j=1, 2, \ldots, r$, $k_1+k_2+\ldots+k_r=n$}.

[Выберем k_1 шаров и поместим в первый ящик, из оставшихся $n-k_1$ шаров $C_{n-k_1}^{k_2}$ способами выберем k_2 шаров и поместим их во второй ящик и т. д. В результате число элементарных событий, составляющих рассматриваемое событие, равно

$$C_n^{k_1}C_{n-k_1}^{k_2}\ldots C_{n-k_1-\ldots-k_{r-2}}^{k_{r-1}}=\frac{n!}{k_1!k_2!\ldots k_r!}.$$

Поделив на общее число элементарных исходов r^n, получим искомую вероятность.]

1.26. Пример (размещение неразличимых шаров по ящикам). Рассмотрим опыт, при котором наудачу выбирается какое-то решение $\omega=(k_1, k_2, \ldots, k_r)$ уравнения в целых неотрицательных числах $k_1+k_2+k_3\ldots+k_r=n$. Чтобы подсчитать число всех элементарных событий ω, сопоставим каждому ω слово длиной $n+r-1$ из символов $*$ и $+$, заменив в сумме $k_1+k_2+\ldots+k_r$ каждое $k_j>0$ на k_j символов $*$ и отбросив все $k_j=0$. Число всех таких слов равно в соответствии с 1.10 $C_{n+r-1}^{r-1}=C_{n+r-1}^n$. Таким образом, в рассматриваемом опыте имеется C_{n+r-1}^{r-1} исходов, которые полагаем равновозможными.

Предположим, что n шаров размещаются по r ящикам так, что любой шар может быть помещен в любой ящик без ограничения на число шаров в ящиках (см. 1.22). Каждое размещение можно характеризовать строчкой $(k_1, k_2, \ldots, k_r)$, где k_i — число шаров в ящике с номером $i=1, 2, \ldots, n$, и указанием номеров k_i шаров, расположенных в i-м ящике, $i=1, 2, \ldots, n$. В соответствии с 1.25 имеется $n!/(k_1!k_2!\ldots k_r!)$ размещений с распределением числа шаров по ящикам, задаваемым строчкой $(k_1, k_2, \ldots, k_r)$. Если отождествить все такие размещения, т. е. не различать размещения, отличающиеся только номерами шаров в ящиках, то придем к опыту с C_{n+r-1}^{r-1} исходами, причем исход $(k_1, k_2, \ldots, k_r)$ имеет вероятность $r^{-n}n!/(k_1!k_2!\ldots k_r!)$. В связи с этим рассмотренный выше опыт с C_{n+r-1}^{r-1} равновозможными исходами $(k_1, k_2, \ldots, k_r)$ называют случайным размещением *неразличимых* шаров. В статистической механике распределение некоторых типов элементарных частиц по ячейкам, на которые разбивается фазовое пространство, соответствует схеме размещения неразличимых шаров.

1.27. Задача. В схеме случайного размещения n неразличимых шаров по r ящикам (см. 1.26) найти вероятности событий: (I) все ящики окажутся занятыми (при $r \leqslant n$); (II) ровно m ящиков окажутся пустыми; (III) заданные k ящиков содержат в совокупности m шаров.

[(I) Число положительных решений уравнения $k_1+k_2+\ldots+k_r=n$ совпадает с числом неотрицательных решений уравнения $l_1+l_2+\ldots+l_r=n-r$, так как можно положить $l_i=k_i-1$, $i=1, \ldots, r$. Отсюда искомая вероятность равна $C_{n-1}^{r-1}/C_{n+r-1}^{r-1}$. Иначе говоря, чтобы получить размещение без пустых ящиков, сначала помещаем в каждый ящик по одному шару, а затем раскладываем оставшиеся $n-r$ шаров по r ящикам любым из $C_{n-r+r-1}^{r-1}=C_{n-1}^{r-1}$ способов. (II) Пустыми можно взять любую из C_r^m групп по m ящиков, в остающиеся $r-m$ ящиков надо разложить n шаров так, чтобы никакой из этих ящиков не был пуст, что возможно сделать C_{n-1}^{r-m-1} способами. Отсюда получаем вероятность $C_r^m C_{n-1}^{r-m-1}/C_{n+r-1}^{r-1}$. (III) В заданные k ящиков поместим m шаров C_{m+k-1}^{k-1} различными способами, остающиеся $n-m$ шаров размещаем в остальные $r-k$ ящиков $C_{n-m+r-k-1}^{r-k-1}$ способами. Перемножив полученные числа и поделив на C_{n+r-1}^{r-1}, получим искомую вероятность.]

1.28. Задача. Производится n-кратный случайный выбор без возвращения из урны с N шарами, из которых M — красные. Показать, что вероятность события $\mathscr{B}$={выборка содержит m красных шаров} равна

$$p_n(m;\ N,\ M)=C_M^m C_{N-M}^{n-m}/C_N^n.$$

[Множество элементарных событий, составляющих $\mathscr{B}$, разобьем на C_n^m одинаковых по численности частей, соответствующих определенному набору номеров испытаний $k_1<k_2<\ldots<k_m$, при которых появились красные шары. Число элементарных событий с заданным набором $(k_1, k_2, \ldots, k_m)$ равно

$$M(M-1)\ldots(M-m+1)(N-M)(N-M-1)\ldots(N-M-(n-m)+1).$$

В самом деле, при заданных номерах $k_1<k_2<\ldots<k_m$ появления

красных шаров имеется взаимное однозначное соответствие между элементарными событиями

$$\omega=(i_1, \ldots, i_{k_1-1}, i_{k_1}, i_{k_1+1}, \ldots, i_{k_2-1}, i_{k_2}, i_{k_2+1}, \ldots, i_n)$$

и парами последовательностей

$$(i_1, \ldots, i_{k_1-1}, i_{k_1+1}, \ldots, i_{k_1-1}, i_{k_2+1}, \ldots, i_n), (i_{k_1}, i_{k_2}, \ldots, i_{k_m}),$$

которые являются выборками без возвращения из множеств из $N-M$ и M элементов соответственно.

Всего имеется C_n^m наборов $(k_1, k_2, \ldots, k_m)$, так что событие $\mathscr{B}$ состоит из

$$C_n^m (M!/(M-m)!)((N-M)!/(N-M-(n-m))!)$$

элементарных событий. Разделив полученное число на число всех элементарных событий опыта

$$N(N-1)\ldots(N-n+1)=N!/(N-n)!,$$

находим вероятность события $\mathscr{B}$

$$\frac{n!M!\,(N-M)!\,(N-n)!}{m!\,(n-m)!\,(M-m)!\,(N-M-(n-m))!\,N!}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}.\ ]$$

1.29. Пример (выбор без возвращения и без учета порядка внутри выборки). Представим себе, что из урны с N занумерованными шарами извлекаются без возвращения $n \leqslant N$ шаров и откладываются без учета порядка появления. Таким образом, исход опыта — подмножество $\{i_1, i_2, \ldots, i_n\}$ из номеров извлеченных шаров. Всего имеется C_N^n подмножеств из n элементов. Полагая все C_N^n элементарных событий равновозможными, приходим к схеме случайного выбора без возвращения и без учета порядка внутри выборки. Эта схема может быть применена и в случае, когда производится выбор с учетом порядка внутри выборки, но интересующее нас событие от этого порядка не зависит.

1.30. Задача. Решить задачу 1.28, используя схему 1.29. [Все подмножества из n шаров, содержащие m красных, можно получить, объединяя любое из C_M^m подмножеств из m красных шаров с любым из C_{N-M}^{n-m} подмножеств из $n-m$ шаров других цветов. Отсюда событие $\mathscr{B}$ состоит из $C_M^m C_{N-M}^{n-m}$ элементарных событий. Поделив это число на число C_N^n всех элементарных событий, получим ответ.]

§ 2.

ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО, СЛУЧАЙНЫЕ ВЕЛИЧИНЫ, РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

Примеры и задачи § 1 наглядно демонстрируют возможности анализа вероятностных опытов за счет их математического описания. Элементарным событиям (исходам) опыта ставятся во взаимно однозначное соответствие точки ω некоторого множества Ω, любое событие рассматривается как логическое объединение элементарных событий и описывается соответствующим подмножеством из Ω. Для опытов с равновозможными исходами вероятность события, описываемого подмножеством $A\subseteq\Omega$, определяется как отношение M_A/N числа M_A элементов в A (т. е. числа элементарных событий, составляющих заданное событие), к числу N элементов в Ω (т. е. полному числу элементарных исходов опыта). Для перехода к формальной математической модели опыта с равновозможными исходами остается сделать последний шаг: вынести соответствие между событиями с одной стороны и подмножествами из Ω — с другой за пределы теории, сохраняя это соответствие для интерпретации математических результатов, и объявить множество Ω с определенной на его подмножествах функцией $\mathbf{P}(A)=M_A/N$ непосредственным объектом теории. Следует при этом подчеркнуть, что содержательная интерпретация вероятностных моделей играет решающую роль в формировании проблематики теории вероятностей и методов анализа. Не случайно в теории вероятностей сохраняется наряду с формальной и содержательная терминология: подмножество $A\subseteq\Omega$ называют также *событием*, точку $\omega\in\Omega$ — *элементарным событием*, вместо выражения «пусть $\omega\in A$» говорят «пусть произошло событие A». Логические операции над событиями как высказываниями переходят в теоретико-множественные операции над событиями как подмножествами: $A\cup B$, $A\cap B(\equiv AB)$, $\bar{A}$, $A\subseteq B$, однако к последним применяется также и логическая терминология: «произошло A или B», «произошли и A, и B», «A не произошло», «A влечет B». Фактически при этом подразумевается умозрительный эксперимент — случайный выбор элемента ω из множества Ω, а событие A воспринимается как утверждение $\omega\in A$.

2.1. Задача. Записать в теоретико-множественных обозначениях события, приведенные в 1.1.

$$[A\bar{B}\bar{C},\ AB\bar{C},\ ABC,\ A\cup B\cup C,\ A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C,\ AB\cup AC\cup BC,$$

$$AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC,\ \bar{A}\bar{B}\bar{C},\ \overline{ABC}.]$$

Вероятность $\mathbf{P}(A)=M_A/N$ была введена в § 1, основываясь на эмпирическом законе устойчивости частот: если опыт, результатом которого может быть событие A, повторяется при идентичных условиях неограниченное число раз, то частота появления события A в первых n испытаниях стремится с ростом n к $\mathbf{P}(A)$. Таким образом, не будучи в состоянии ответить на вопрос, произойдет событие A в каждом отдельном испытании или нет, мы измеряем возможность осуществления события A предельным значением его частоты в длинном ряду испытаний. Несмотря на всю важность частотной интерпретации вероятности, следует подчеркнуть, что вероятность как числовая мера возможности осуществления события имеет свой внутренний, интуитивно прозрачный смысл, вытекающий из непосредственных представлений об опыте со случайными исходами. Так, сказав, что некоторое событие A имеет вероятность 3/5, можно представить урну с 5 шарами, из которых 3 красных. Тогда событие «при случайном выборе получен красный шар» с вероятностной точки зрения эквивалентно A. Поясним, почему выбор числовой меры в виде $\mathbf{P}(A)=M_A/N$ является достаточно естественным. Рассмотрим для примера опыт с извлечением наудачу игральной карты из колоды. Без всякой теории ясно, что события {получен туз}, {получен король}, {получена дама} и т. д., или события {получена пика}, {получена трефа} и т. д. имеют одинаковые возможности осуществления. Это означает, что любая числовая мера, претендующая на роль измерителя возможности осуществления события в рассматриваемой схеме, должна зависеть только от числа M элементарных событий, составляющих заданное событие, и числа N всех исходов (карт в колоде). Складывая вместе несколько одинаковых колод, мы, очевидно, не изменим шансов наступления событий {получен туз}, {получена пика} и т. д., так что числовая мера должна зависеть от M и N через их отношение. Выбор в качестве такой меры самого отношения M/N имеет несомненные преимущества еще и потому, что мера $\mathbf{P}(A)$ обладает весьма привлекательным свойством аддитивности

$$\mathbf{P}(A \cup B)=\mathbf{P}(A)+\mathbf{P}(B), \text{ если } AB=\varnothing.$$

Это свойство объединяет понятие вероятности с понятиями физической массы, длины, площади и др.

Вероятностную меру доопределяют на пустом подмножестве, полагая $\mathbf{P}(\varnothing)=0$, что дает значительные технические удобства.

2.2. Задача. Из аддитивности вероятностной меры $\mathbf{P}(A)$ вывести следующие ее свойства:

(I) $\mathbf{P}(\bar{A})=1-\mathbf{P}(A)$;

(II) если $B \subseteq A$, то $\mathbf{P}(A \ominus B)=\mathbf{P}(A)-\mathbf{P}(B)$, $A \ominus B \equiv A\bar{B}$;

(III) если $B \subseteq A$, то $\mathbf{P}(B) \leqslant \mathbf{P}(A)$;

(IV) $\mathbf{P}(A \cup B)=\mathbf{P}(A)+\mathbf{P}(B)-\mathbf{P}(A\bar{B})$;

(V) $\mathbf{P}(A \cup B) \leqslant \mathbf{P}(A)+\mathbf{P}(B)$;

(VI) $\mathbf{P}(A \cup B \cup C)=\mathbf{P}(A)+\mathbf{P}(B)+\mathbf{P}(C)-\mathbf{P}(AB)-\mathbf{P}(AC)-$
$-\mathbf{P}(BC)+\mathbf{P}(ABC)$.

[(I) $1=\mathbf{P}(\Omega)=\mathbf{P}(A \cup \bar{A})=\mathbf{P}(A)+\mathbf{P}(\bar{A})$, так как $A\bar{A}=\varnothing$. (II) $\mathbf{P}(A)=$ $=\mathbf{P}(A\bar{B} \cup B)=\mathbf{P}(A\bar{B})+\mathbf{P}(B)$, так как $A\bar{B}B=\varnothing$. (IV) Представив событие $A \cup B$ в виде объединения трех непересекающихся событий $A\overline{AB}$, AB, $B\overline{AB}$ и воспользовавшись свойством (II), получаем

$$\mathbf{P}(A \cup B) = \mathbf{P}(A\overline{AB}) + \mathbf{P}(AB) + \mathbf{P}(B\overline{AB}) = \mathbf{P}(A) - \mathbf{P}(AB) + \\ + \mathbf{P}(AB) + \mathbf{P}(B) - \mathbf{P}(AB) = \mathbf{P}(A) + \mathbf{P}(B) - P(AB).$$

(VI) $\mathbf{P}(A \cup B \cup C) = \mathbf{P}((A \cup B) \cup C) = \mathbf{P}(A \cup B) + P(C) - \mathbf{P}((A \cup B)C) = \\ = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(AB) + \mathbf{P}(C) - \mathbf{P}(AC) - \mathbf{P}(BC) + \mathbf{P}(ABC)$.]

На рис. 10, 11, иллюстрирующих формулы (IV), (VI), множества A, B, C представлены кружками, цифрами указана кратность, с которой формула под рисунком учитывает меру каждой области.

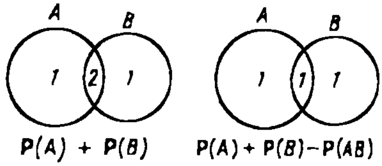

Рис. 10

2.3. Задача. Решить задачу 1.2 с помощью формулы 2.2 (VI).

[Пусть A_i — событие «i-е письмо попало в свой конверт», $i = 1, 2, 3$. Имеется два разложения писем по конвертам, отвечающие событию A_i. Событие A_iA_j, $i \neq j$, совпадает с $A_1A_2A_3$, и ему соответствует одно разложение. Учитывая, что всего имеется $3! = 6$ разложений, получаем $\mathbf{P}(A_1 \cup A_2 \cup A_3) = 3 \cdot 2/6 - 3 \cdot 1/6 + 1/6 = 4/6$.]

Опыт с равновозможными исходами представляет собой опре-

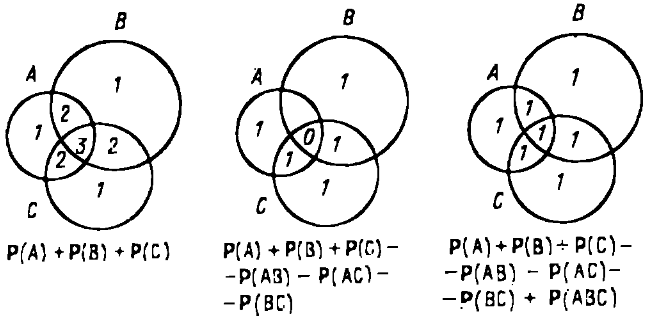

Рис. 11

деленную идеализацию действительности. Любая монета, игральная кость не абсолютно симметричны, и, проводя очень большое число испытаний, обнаружим, что частоты исходов отклоняются от N^{-1} больше, чем это следовало бы ожидать в предположении равновозможности всех N исходов. Эмпирический закон устойчивости частот тем не менее выполняется, однако предельные частоты p_i, $i = 1, 2, \ldots, N$ исходов оказываются некоторыми числами, вообще говоря отличными от N^{-1}. Их и следует принять за вероятности элементарных событий. Учитывая, что предельная частота появления любого события равна сумме частот составляющих его элементарных событий, приходим к следующей математической модели опыта, исходы которого, вообще говоря, не равновозможны.

Пусть $\Omega=\{\omega\}$ — произвольное конечное или счетное множество, которое назовем *пространством элементарных событий*, $p(\omega)$ — числовая функция на Ω, $0\leqslant p(\omega)\leqslant 1$, и

$$\Sigma_{\omega\in\Omega}p(\omega)=1,$$

которую назовем *распределением вероятностей на* Ω. *Вероятность* любого события $A\subseteq\Omega$ определим формулой

$$\mathbf{P}(A)=\sum_{\omega\in A}p(\omega),\ \mathbf{P}(\varnothing)=0.$$

Отметим, что в случае бесконечного Ω сумма бесконечного числа слагаемых понимается как ряд, причем ввиду неотрицательности слагаемых способ упорядочивания членов ряда несуществен.

Если Ω взято конечным, а распределение вероятностей равномерно ($p(\omega)=\text{const}$), то получаем модель опыта с равновозможными исходами. В общем случае вероятностная модель $(\Omega, p(\omega))$ называется *дискретным вероятностным пространством*. Функция $\mathbf{P}(A)$, определенная на подмножествах A, называется *вероятностной мерой*. Она, очевидно, аддитивна

$$\mathbf{P}(A\cup B)=\mathbf{P}(A)+\mathbf{P}(B),\ \text{если}\ AB=\varnothing,$$

и, следовательно, обладает свойствами, которые в задаче 2.2 приведены для случая равномерного распределения вероятностей. Для бесконечного Ω добавляется расширенное свойство аддитивности, называемое *счетной аддитивностью*:

$$\mathbf{P}\Big(\bigcup_{i=1}^{\infty}A_i\Big)=\sum_{i=1}^{\infty}\mathbf{P}(A_i),\ \text{если}\ A_iA_j=\varnothing\ \text{при}\ i\neq j.$$

В самом деле,

$$\sum_{\omega\in\cup A_i}p(\omega)=\sum_{(i)}\sum_{\omega\in A_i}p(\omega)=\sum_{(i)}\mathbf{P}(A_i),$$

где перестановка членов ряда законна в силу их неотрицательности.

Пространство $(\Omega, p(\omega))$ с равномерным распределением вероятностей $p(\omega)=\text{const}$ может служить источником многих интересных примеров дискретных вероятностных пространств за счет «склеивания» элементарных событий ω. Так, если из N шаров в урновой модели (см. 1.20) M_i шаров окрашены в i-й цвет, $i=1, 2, \ldots, k$, $M_1+M_2+\ldots+M_k=N$, то, обращая внимание лишь на цвет извлеченного шара, приходим к опыту с k исходами и вероятностями исходов $p_i=M_i/N$, $i=1, \ldots, k$. В общем случае пусть

B_1, B_2, ... — произвольная совокупность попарно непересекающихся событий $B_i \in \Omega$, в сумме составляющих все Ω:

$$\Omega = \bigcup_{(i)} B_i, \quad B_i B_j = \varnothing \text{ при } i \neq j.$$

Назовем систему событий B_1, B_2, ... *разбиением пространства* Ω или *полной группой событий*. Введем отображение $X(\omega)$ пространства Ω на некоторое множество $\mathscr{X} = \{x_1, x_2, \ldots\}$, при котором все $\omega \in B_i$ переходят в x_i: $X(\omega) = x_i$ при $\omega \in B_i$, $i = 1, 2, \ldots$. Введем на $\mathscr{X}$ числовую функцию $p_X(x_i) = \mathbf{P}(B_i)$. Так как

$$\sum_{x \in \mathscr{X}} p_X(x) = \sum_{(i)} \mathbf{P}(B_i) = \mathbf{P}(\bigcup_{(i)} B_i) = \mathbf{P}(\Omega) = 1,$$

то $p_X(x)$ — распределение вероятностей на $\mathscr{X}$. Дискретное вероятностное пространство $(\mathscr{X}, p_X(x))$ называют *индуцированным* (*порожденным*) *отображением* X.

2.4. Пример (биномиальное распределение). В модели n-кратного случайного выбора с возвращением из урны с долей p красных шаров рассмотрим полную группу событий $B_k = \{$выборка содержит k красных шаров$\}$, $k = 0, 1, \ldots, n$. Отображение $X(\omega) = k$, $\omega \in B_k$, $k = 0, 1, \ldots, n$, порождает распределение вероятностей на множестве $\mathscr{X} = \{0, 1, \ldots, n\}$ (см. 1.21):

$$p_X(k) = p_n(k, p) = C_n^k p^k (1-p)^{n-k}, \; k = 0, 1, \ldots, n.$$

Отсюда следует, что

$$\sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = 1.$$

при любом рациональном p, $0 < p < 1$. Выбрав последовательность рациональных чисел, сходящуюся к заданному действительному числу $0 < p < 1$, получаем, что предыдущее равенство справедливо при любом действительном $0 < p < 1$. Это значит, что на множестве $\mathscr{X}$ функция $p_n(k, p)$ задает распределение вероятностей при любом $0 < p < 1$. Распределение $p_n(k, p)$ называется *биномиальным с параметрами n и p*.

2.5. Пример (испытания Бернулли). Пространство Ω в 2.4 отобразим на множество $\mathscr{X}_n = \{(x_1, \ldots, x_n)\}$ всех последовательностей из 0 и 1, полагая $x_i = 1$, если i-й шар красный, и $x_i = 0$ — в противном случае. Тогда (см. 1.21)

$$p_X(x) = p^{s_n} (1-p)^{n-s_n}, \; s_n = x_1 + \ldots + x_n.$$

При этом

$$\sum_{x \in \mathscr{X}_n} p^{s_n} (1-p)^{n-s_n} = \sum_{m=0}^{n} C_n^m p^m (1-p)^{n-m} = 1.$$

Предположим теперь, что $0<p<1$ — произвольное действительное число. Функция

$$p(x)=p^{s_n}(1-p)^{n-s_n}, \quad s_n=x_1+\ldots+x_n,$$

как вытекает из 2.4, задает на $\mathscr{X}_n$ распределение вероятностей. Дискретное вероятностное пространство $(\mathscr{X}_n, p(x))$ называют *n испытаниями Бернулли с вероятностью успеха p*. В урновой интерпретации успех — появление красного шара.

Отображение $Y(x)=x_1+\ldots+x_n$ пространства $\mathscr{X}_n$ на множество $\mathscr{Y}=\{0, 1, \ldots, n\}$ порождает на $\mathscr{Y}$ биномиальное распределение

$$p_Y(k)=C_n^k p^k(1-p)^{n-k}, \quad k=0, 1, \ldots, n.$$

2.6. З а д а ч а. Рассмотрим отображение пространства $(\mathscr{X}_n, p(x))$ испытаний Бернулли на множество $\mathscr{X}_m=\{0, 1\}^m$, $m<n$: $(x_1, \ldots, x_n)\to(x_1, \ldots, x_m)$. Показать, что это отображение порождает m испытаний Бернулли. Аналогичное утверждение справедливо и для отображения $(x_1, \ldots, x_n)\to(x_{i_1}, x_{i_2}, \ldots, x_{i_m})$, где $1\leqslant i_1<i_2<\ldots<i_m\leqslant n$ фиксированы.

[В урновой интерпретации сформулированное утверждение означает, что если при случайном n-кратном выборе с возвращением отмечать лишь результаты испытаний с номерами i_1, $i_2, \ldots$ $\ldots, i_m$, то получим случайную выборку с возвращением объема m. Утверждение представляется очевидным, но для иррационального p все же требует доказательства. Фиксируем произвольную последовательность $y=(y_1, \ldots, y_m)$ из 0 и 1. Заметив, что $p(x)=p(x_1, \ldots, x_n)=p(x_1, \ldots, x_m)\cdot p(x_{m+1}, \ldots, x_n)$, найдем вероятность события $B_y=\{x: x_1=y_1, \ldots, x_n=y_m\}$, что первые m испытаний привели к результату $y=(y_1, \ldots, y_m)$:

$$\mathbf{P}(B_y)=\sum_{x\in B_y} p(x_1, \ldots, x_m)p(x_{m+1}, \ldots, x_n)=p(y_1, \ldots, y_m)\times$$

$$\times\sum_{(x_{m+1}, \ldots, x_n)\in\mathscr{X}_{n-m}} p(x_{m+1}, \ldots, x_n)=p(y_1, \ldots, y_m).]$$

Функции, отображающие дискретное вероятностное пространство $(\Omega, p(\omega))$ в числовую прямую R, называют *дискретными случайными величинами* (сл. в.). Пусть $X(\omega)$ — некоторая случайная величина, $\mathscr{X}$ — множество значений X, $p_X(x)$ — распределение вероятностей на $\mathscr{X}$, задаваемое формулой

$$p_X(x)=\sum_{\{\omega: X(\omega)=x\}} p(\omega)=\mathbf{P}(\omega: X(\omega)=x).$$

Распределение $p_X(x)$ называют *распределением вероятностей сл. в. X*; при фиксированном x $p_X(x)$ называют *вероятностью значения x сл. в. X*.

В значительной части примеров и задач § 1 на самом деле вычислялось распределение вероятностей различных случайных величин. Так, число проб, пока не будет найден нужный ключ в 1.6 (I), имеет *равномерное* распределение вероятностей: $p(x)=n^{-1}$ при $x=1, 2, \ldots, n$; число человек между A и B при случайной расстановке в ряд n человек в 1.7 имеет распределение вероятностей $p(x)=2(1-x/(n-1))n^{-1}$, $x=0, 1, \ldots, n-2$; сл. в. S_n — положение частицы в момент времени n в симметричном случайном блуждании — имеет распределение $p(k)=C_n^{(n+k)/2}2^{-n}$, $|k|\leqslant n$, k одной четности с n.

Сл. в. X с распределением вероятностей $p_X(x)$ порождает дискретное вероятностное пространство $(\mathscr{X}, p_X(x))$. Наоборот, если имеется произвольное дискретное вероятностное пространство $(\mathscr{X}, p(x))$, где $\mathscr{X}\subseteq R$, то можно построить сл. в. X, для которой $p_X(x)=p(x)$. Именно, определим сл. в. X как функцию на $\mathscr{X}$, полагая $X(x)=x$, так что событие {сл. в. X приняла значение x} — элементарное событие вероятностного пространства $(\mathscr{X}, p(x))$ и имеет вероятность $p(x)$. Так определенная случайная величина называется *непосредственно заданной*. Отмеченное обстоятельство позволяет говорить о случайной величине с заданным распределением вероятностей безотносительно к вероятностному пространству, на котором она определена, поскольку можно считать, что случайная величина непосредственно задана. В определенном смысле случайная величина и порожденное ею вероятностное пространство — эквивалентные формы представления вероятностного опыта.

Предположим, что на вероятностном пространстве $\{\Omega, p(\omega)\}$ рассматривается набор сл. в. $\mathbf{X}(\omega)=(X_1(\omega), \ldots, X_n(\omega))$. На множестве $\mathscr{X}=\mathscr{X}_1\times\mathscr{X}_2\times\ldots\times\mathscr{X}_n$ значений *случайного вектора* $\mathbf{X}(\omega)$ отображение $\mathbf{X}(\omega)$ индуцирует вероятностное распределение $p_\mathbf{X}(\mathbf{x})$, которое называют *совместным распределением вероятностей* случайных величин $X_1, \ldots, X_n$. Как и в одномерном случае, любое распределение вероятностей $p(\mathbf{x})$ на $\mathscr{X}=\mathscr{X}_1\times\ldots\times\mathscr{X}_n$ позволяет определить непосредственно набор случайных величин $X_1(\mathbf{x})=x_1, \ldots, X_n(\mathbf{x})=x_n$, $\mathbf{x}=(x_1, \ldots, x_n)$, с совместным распределением вероятностей $p(\mathbf{x})$. Как пример приведем пару случайных величин: M_n — максимум траектории симметричного случайного блуждания длительности n, S_n — положение блуждающей частицы в конечный момент времени. Их совместное распределение вероятностей (см. 1.17)

$$p_{M_n, S_n}(r, k)=2^{-n}\left(C_n^{r+(n-k)/2}-C_n^{r+1+(n-k)/2}\right),$$

где C_n^j полагается равным нулю при $j<0$, $j>n$ и j не целом.

В качестве другого примера рассмотрим набор случайных величин $\mathbf{X}=(X_1, \ldots, X_n)$, принимающих значения 0 и 1, которые зададим совместным распределением

$$p_\mathbf{X}(\mathbf{x})=p^{s_n}(1-p)^{n-s_n},\ \mathbf{x}=(x_1, \ldots, x_n),\ s_n=x_1+\ldots+x_n.$$

Набор случайных величин $\mathbf{X}$ порождает n испытаний Бернулли с вероятностью успеха (выпадения 1), равной p; для последовательности сл. в. $X_1, X_2, \ldots, X_n$ сохраняется название испытания Бернулли.

2.7. Задача. Показать, что для любого разбиения $B_1, B_2, \ldots$ вероятностного пространства и любого события A справедлива формула

$$\mathbf{P}(A)=\sum_{\{k\}}\mathbf{P}(AB_k).$$

[Пересекая обе части равенства $\Omega=\bigcup_{\{k\}}B_k$ с событием A, получаем $A=\bigcup_{\{k\}}AB_k$. Остается воспользоваться (счетной) аддитивностью вероятностной меры $\mathbf{P}$.]

2.8. Задача. Известно совместное распределение $p_{X,Y}(x, y)$ пары случайных величин X, Y. Показать, что частные распределения $p_X(x)$ и $p_Y(y)$ выражаются через их совместное формулами

$$p_X(x)=\sum_{\{y\}}p_{X,Y}(x, y), \quad p_Y(y)=\sum_{\{x\}}p_{X,Y}(x, y),$$

где суммирование ведется по множествам $\mathscr{Y}$, $\mathscr{X}$ значений сл. в. Y и сл. в. X соответственно. Обобщить утверждение на произвольное число случайных величин.

[События $A_x=\{\omega:X(\omega)=x\}$, $x\in\mathscr{X}$, образуют полную группу событий. Это же верно относительно событий $B_y=\{\omega:Y(\omega)=y\}$, $y\in\mathscr{Y}$. Отсюда, применяя 2.7, получаем

$$\mathbf{P}(A_x)=\sum_{\{y\}}\mathbf{P}(A_xB_y), \quad \mathbf{P}(B_y)=\sum_{\{x\}}\mathbf{P}(B_yA_x).$$

Остается заметить, что

$$\mathbf{P}(A_xB_y)=\mathbf{P}(\omega:X(\omega)=x,\ Y(\omega)=y)=p_{X,Y}(x, y).$$

Для n случайных величин $X_1, \ldots, X_n$ и любого набора индексов $1\leqslant i_1<i_2<\ldots<i_m\leqslant n$ имеем аналогично

$$p_{X_{i(1)},\ldots,X_{i(m)}}(x_{i_1}, \ldots, x_{i_m})=\sum_{\{x_{j(1)},\ldots,x_{j(n-m)}\}}p_{X_1,\ldots,X_n}(x_1, \ldots, x_n),$$

где $\{j_1, j_2, \ldots, j_{n-m}\}$ — дополнительное к $\{i_1, i_2, \ldots, i_m\}$ множество индексов, а суммирование ведется по всем возможным значениям $x_{j_1}, \ldots, x_{j_{n-m}}$, $i(k)\equiv i_k$, $j(k)\equiv j_k$.]

2.9. Задача. Пусть $X_1, \ldots, X_n$ — испытания Бернулли с вероятностью успеха p. Найти частные распределения сл. в. X_k, $k=1, \ldots, n$.

[Используя 2.8 и 2.6, находим $p_{X_k}(x_k)=p^{x_k}(1-p)^{1-x_k}$, $x_k=0, 1$. Отметим равенство

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)=p_{X_1}(x_1)\ldots p_{X_n}(x_n).]$$

2.10. Задача. Проводится n испытаний Бернулли. Обозначим U число успехов в первых m испытаниях и V — число успехов в последующих $n-m$ испытаниях. Найти совместное распределение вероятностей сл. в. U, V.

[Пусть $(\mathscr{X}, p(x))$ — n испытаний Бернулли с вероятностью успеха $p: \mathbf{x}=(x_1, \ldots, x_n)$, $x_i=0$ или 1, $i=1, \ldots, n$, $s_k=x_1+\ldots+x_k$, $k=1, \ldots, n$, $p(x)=p^{s_n}(1-p)^{n-s_n}$, $U(x)=s_m$, $V(x)=s_n-s_m$. Используя равенство $p(x)=p(x_1, \ldots, x_n)=p(x_1, \ldots, x_m)\,p(x_{m+1}, \ldots, x_n)$, получаем

$$p_{U,V}(u, v)=\sum_{\{x:s_m=u,\,s_n-s_m=v\}} p(x_1, \ldots, x_m)\,p(x_{m+1}, \ldots, x_n)=$$

$$=\sum_{\{x_1,\ldots,x_m:s_m=u\}} p(x_1, \ldots, x_m) \sum_{\{x_{m+1},\ldots,x_n:s_n-s_m=v\}} p(x_{m+1}, \ldots, x_n)=$$

$$=C_m^u p^u (1-p)^{m-u}\, C_{n-m}^v p^v (1-p)^{n-m-v}.$$

Суммируя $p_{U,V}(u, v)$ по u и v, находим частные распределения вероятностей сл. в. U и V:

$$p_U(u)=C_m^u p^u (1-p)^{m-u}, \quad p_V(v)=C_{n-m}^v p^v (1-p)^{n-m-v}.$$

Отметим равенство

$$p_{U,V}(u, v)=p_U(u)\,p_V(v).]$$

2.11. Пример (простое случайное блуждание). Пусть $X_1, X_2, \ldots, X_n$ — последовательность испытаний Бернулли с вероятностью успеха p. Введем сл. в. $Y_k=2X_k-1$, $k=1, 2, \ldots, n$. Сл. в. Y_k принимают значения ± 1 и

$$p_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)=p_{X_1,\ldots,X_n}((y_1+1)/2, \ldots, (y_n+1)/2)=$$

$$=p^{(n+y_1+\ldots+y_n)/2}(1-p)^{(n-y_1-\ldots-y_n)/2}.$$

Положим $Z_k=Y_1+\ldots+Y_k$, $k=1, \ldots, n$, $Z_0=0$. Назовем последовательность сл. в. $Z_0, Z_1, \ldots, Z_n$ *простым случайным блужданием*, начинающимся из нуля. При $p=1/2$ рассматриваемая модель совпадает с симметричным случайным блужданием, изучавшимся в § 1.

Расчеты, проведенные для симметричного блуждания, оказываются полезными при вычислении вероятностей для произвольного p, если заметить, что все траектории $(Z_0, Z_1, \ldots, Z_n)$ с фиксированным значением Z_n имеют одинаковые вероятности:

$$p_{Z_1,\ldots,Z_n}(j_1, \ldots, j_n)=p_{Y_1,Y_2,\ldots,Y_n}(j_1, j_2-j_1, \ldots, j_n-j_{n-1})=$$

$$=p^{(n+j_n)/2}(1-p)^{(n-j_n)/2},$$

если $j_1, j_2-j_1, \ldots, j_n-j_{n-1}$ принимают значения ± 1, и $p_{Z_1,\ldots,Z_n}(j_1, \ldots, j_n)=0$ — в противном случае. Вычислим, например, распределение вероятностей конечного положения Z_n блуждающей частицы. Используя 2.8 получаем

$$p_{Z_n}(j_n) = \sum_{\{j_1,\ldots,j_{n-1}\}} p^{(n+j_n)/2}(1-p)^{(n-j_n)/2},$$

где суммирование ведется по всем наборам $j_1, \ldots, j_{n-1}$, таким, что $(0, j_1, j_2, \ldots, j_{n-1}\ j_n)$ представляет собой траекторию блуждания, т. е. $j_1, j_2-j_1, \ldots, j_n-j_{n-1}$ равны 1 или —1. Число всех таких траекторий найдено в 1.12, откуда получаем

$$p_{Z_n}(k) = C_n^{(n+k)/2} p^{(n+k)/2}(1-p)^{(n-k)/2}$$

при $|k| \leqslant n$, k одной четности с n.

2.12. Задача. Найти вероятность, что блуждающая частица модели 2.11 в течение $2n$ единиц времени будет находиться на положительной полуоси.

[Вероятность события $A = \{\omega : Z_1(\omega) > 0, \ldots, Z_{2n}(\omega) > 0\}$ запишем по формуле 2.7 через полную группу событий

$$B_k = \{\omega : Z_{2n}(\omega) = 2k\}, \quad k = 0, \pm 1, \ldots, \pm n:$$

$$\mathbf{P}(A) = \sum_{k=1}^{n} \mathbf{P}(\omega : Z_1(\omega) > 0, \ldots, Z_{2n-1}(\omega) > 0, \quad Z_{2n}(\omega) = 2k) =$$

$$= \sum_{k=1}^{n} \sum_{\{j_1>0,\ldots,j_{2n-1}>0\}} p^{n+k}(1-p)^{n-k},$$

где внутренняя сумма берется по всем наборам $j_1 > 0, \ldots, j_{2n-1} > 0$, таким, что $(0, j_1, \ldots, j_{2n-1}, 2k)$ представляет собой траекторию блуждания. Число указанных траекторий равно (см. 1.14)

$$C_{2n-1}^{n+k-1} - C_{2n-1}^{n+k} = \frac{(2n-1)!}{(n+k-1)!\,(n-k-1)!}\left(\frac{1}{n-k} - \frac{1}{n+k}\right) = \frac{k}{n} C_{2n}^{n+k},$$

так что

$$\mathbf{P}(A) = \sum_{k=1}^{n} (k/n) C_{2n}^{n+k} p^{n+k}(1-p)^{n-k}.]$$

2.13. Задача. Найти распределение вероятностей сл. в. $M_n = \max(Z_0, Z_1, \ldots, Z_n)$ — максимума траектории простого случайного блуждания 2.11.

[Проводя вычисления по схеме 2.12 и используя 1.17, находим

$$\mathbf{P}(\omega : M_n(\omega) = k) = \sum_{k=-n}^{r} \mathbf{P}(\omega : M_n(\omega) = r, \ Z_n(\omega) = k) =$$

$$= \sum_{k=-n}^{r} \sum_{\{\max(j_1,\ldots,j_{n-1},k)=r\}} \mathbf{P}(\omega : Z_1(\omega) = j_1, \ldots, Z_{n-1}(\omega) = j_{n-1}, Z_n(\omega) = k) =$$

$$= \sum_{k=-n+2r}^{r} \left(C_n^{r+(n-k)/2} - C_n^{r+1+(n-k)/2}\right) p^{(n+k)/2} (1-p)^{(n-k)/2} =$$

$$= \sum_{k=-n+2r}^{r} \frac{2r-k+1}{n+1} C_{n+1}^{r+1+(n-k)/2} p^{(n+k)/2} (1-p)^{(n-k)/2}$$

при $r=0, 1, \ldots, n$, где $C_m^s=0$ при $s<0$, $s>m$ и s не целом.]

2.14. Пример (геометрическое распределение). Рассмотрим испытания Бернулли $X_1, X_2, \ldots$ с вероятностью успеха p. Обозначим T_1 число испытаний, предшествующих первому успеху. Если общее число испытаний равно n, то $T_1=k \Leftrightarrow X_1=\ldots=X_k=0$, $X_{k+1}=1$, $k \leqslant n-1$, и T_1 не определено, когда $X_1=\ldots=X_n=0$, так что

$$\mathbf{P}(\omega : T_1(\omega)=k)=p_{X_1,\ldots,X_{k+1}}(0, \ldots, 0, 1)=(1-p)^k p, \quad k=0, 1, \ldots, n-1$$

и $T_1(\omega)$ не определена с вероятностью $(1-p)^n$. Говорят, что $T_1(\omega)$ является *дефектной* случайной величиной, «дефект» распределения вероятностей $p_{T_1}(k)$, $k=0, 1, \ldots, n-1$, равен $1-p_{T_1}(0)-$ $-p_{T_1}(1)-\ldots-p_{T_1}(n-1)=(1-p)^n$. Устремляя n к ∞, приходим к распределению вероятностей

$$p(k)=(1-p)^k p, \quad k=0, 1, 2, \ldots,$$

сосредоточенному в целых неотрицательных точках прямой, это распределение называется *геометрическим*.

Естественной областью определения сл. в. T_1 могло бы быть вероятностное пространство, описывающее бесконечную последовательность испытаний Бернулли. За элементарные события такого опыта можно принять бесконечные последовательности $\omega=$ $=(x_1, x_2 \ldots)$ из 0 и 1 и положить $T_1(\omega)=\min\{k : x_{k+1}=1\}$. За исключением последовательности ω, целиком состоящей из нулей, $T_1(\omega)$ определена при всех ω. Затруднение, однако, возникает в связи с тем, что множество $\Omega=\{\omega\}$ — несчетное, и определение вероятностной меры на подмножествах Ω требует более серьезных усилий.

2.15. Пример (бесконечная последовательность испытаний Бернулли). Опыт по выбору с возвращением потенциально допускает неограниченное продолжение. Математический подход к анализу неограниченно продолжающихся процессов состоит в представлении бесконечности как завершенного целого. Так, бесконечную последовательность $\mathbf{x}=(x_1, x_2, \ldots)$ из 0 и 1 мы рассматриваем как элементарное событие неограниченно продолжающихся испытаний Бернулли. Обозначим $\mathscr{X}=\{\mathbf{x}\}$ множество всех таких последовательностей, положим $[\mathbf{x}]_n=(x_1, x_2, \ldots, x_n)$, $\mathbf{x}=$ $=(x_1, x_2, \ldots)$, $\mathscr{X}_n=\{(x_1, x_2, \ldots, x_n)\}$. Предположим, что на подмножествах пространства $\mathscr{X}$ определена вероятностная мера $\mathbf{P}(A)$, приводящая к вероятностной модели неограниченной последовательности испытаний Бернулли с вероятностью успеха выпадения 1), равной p. В таком случае множество $\mathscr{X}_n$ начальных отрезков длины n последовательности неограниченных испытаний должно рассматриваться как модель n испытаний Бернулли. Выражаясь более точно, отображение $\mathbf{x} \to [\mathbf{x}]_n$ множества $\mathscr{X}$ в $\mathscr{X}_n$ должно порождать дискретное вероятностное пространство $(\mathscr{X}_n, p(x_1, \ldots \ldots, x_n))$, $p(x_1, \ldots, x_n) = p^{s_n}(1-p)^{n-s_n}$, $s_n = x_1 + \ldots + x_n$. Таким образом, приходим к равенствам

$$\mathbf{P}(\mathbf{x} \in \mathscr{X} : [\mathbf{x}]_n = (x_1^0, \ldots, x_n^0)) = p(x_1^0, \ldots, x_n^0),$$

которые можно положить в основу определения вероятностной меры $\mathbf{P}$ на $\mathscr{X}$.

Введем при каждом n систему событий $\mathscr{A}_n$ пространства $\mathscr{X}$, определяемых результатами первых n испытаний: $A \subseteq \mathscr{X}$ принадлежит системе $\mathscr{A}_n \Leftrightarrow$ вместе с каждым $\mathbf{x} \in A$ в A входят всевозможные $\mathbf{y} \in \mathscr{X}$, такие, что $[\mathbf{y}]_n = [\mathbf{x}]_n$. Иначе говоря, каждое $A \in \mathscr{A}_n$ допускает представление

$$A = \{\mathbf{x} : [\mathbf{x}]_n \in A^{(n)}\}, \quad A^{(n)} \subseteq \mathscr{X}_n,$$

где множество $A^{(n)}$ по заданным A и n определяется однозначно. Легко видеть, что $\mathscr{A}_1 \subseteq \mathscr{A}_2 \subseteq \ldots \subseteq \mathscr{A}_n \subseteq \mathscr{A}_{n+1}$. Положим $\mathscr{A} = \bigcup_{n=1}^{\infty} \mathscr{A}_n$. Зададим на системе событий $\mathscr{A}$ вероятностную меру, полагая

$$\mathbf{P}(A) = \sum_{(x_1, \ldots, x_n) \in A^{(n)}} p(x_1, \ldots, x_n).$$

Это определение корректно: пусть

$$A = \{\mathbf{x} : [\mathbf{x}]_n \in A^{(n)}\} = \{\mathbf{x} : [\mathbf{x}]_m \in A^{(m)}\}, \quad n < m,$$

тогда

$$A^{(m)} = \{(x_1, \ldots, x_m) : (x_1, \ldots, x_n) \in A^{(n)}\}$$

и потому (см. 2.6)

$$\sum_{(x_1, \ldots, x_m) \in A^{(m)}} p(x_1, \ldots, x_m) = \sum_{(x_1, \ldots, x_n) \in A^{(n)}} p(x_1, \ldots, x_n).$$

Класс подмножеств $\mathscr{A}$ недостаточно широк, чтобы выразить многие интересные свойства бесконечной последовательности испытаний. Любое событие $A \in \mathscr{A}$ отражает какое-то свойство лишь некоторой конечной последовательности испытаний. Об определении вероятностей других подмножеств из $\mathscr{X}$ речь будет идти ниже.

2.16. Задача. Показать, что класс множеств $\mathscr{A}$, определенный в 2.14, — *алгебра множеств*, т. е. содержит $\varnothing$, $\mathscr{X}$, вместе с каждым $A \in \mathscr{A}$ содержит дополнение $\bar{A}$, вместе с любыми A, $B \in \mathscr{A}$ — их объединение $A \cup B$. Проверить, что вероятностная функция $\mathbf{P}(A)$, $A \in \mathscr{A}$, введенная в 2.15, аддитивна: $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$, если $AB = \varnothing$.

[Пусть $A = \{\mathbf{x} : [\mathbf{x}]_n \in A^{(n)}\}$, $A_n \subseteq \mathscr{X}_n$, тогда $\bar{A} = \{\mathbf{x} : [\mathbf{x}]_n \in \bar{A}^{(n)}\}$, где $\bar{A}^{(n)}$ — дополнение $A^{(n)}$ в $\mathscr{X}_n$, и, следовательно, $\bar{A} \in \mathscr{A}$. Пусть A, $B \in \mathscr{A}$ и

$$A = \{\mathbf{x} : [\mathbf{x}]_n \in A^{(n)}\}, \quad B = \{\mathbf{x} : [\mathbf{x}]_m \in B^{(m)}\}.$$

Предположим, что $m \geqslant n$. Тогда, полагая

$$A^{(m)} = \{(x_1, \ldots, x_m) : (x_1, \ldots, x_n) \in A^{(n)}\},$$

запишем A в виде

$$A = \{\mathbf{x} : [\mathbf{x}]_m \in A^{(m)}\},$$

откуда

$$A \cup B = \{\mathbf{x} : [\mathbf{x}]_m \in A^{(m)} \cup B^{(m)}\} \in \mathscr{A}.$$

Если $AB = \varnothing$, то $A^{(m)}B^{(m)} = \varnothing$ и

$$\mathbf{P}(A \cup B) = \sum_{(x_1, \ldots, x_m) \in A^{(m)} \cup B^{(m)}} p(x_1, \ldots, x_m) =$$

$$= \sum_{(x_1, \ldots, x_m) \in A^{(m)}} p(x_1, \ldots x_m) + \sum_{(x_1, \ldots, x_m) \in B^{(m)}} p(x_1, \ldots, x_m) =$$

$$= \mathbf{P}(A) + \mathbf{P}(B).]$$

2.17. Задача. В схеме бесконечной последовательности испытаний Бернулли введем сл. в T_k — число неудач между появлениями $(k-1)$-го и k-го успехов, $k = 1, 2, \ldots$, $S_r = T_1 + \ldots + T_r$ — число неудач, предшествовавших r-му успеху. Найти совместное распределение сл. в. $T_1, \ldots, T_r$, их частные распределения и распределение их суммы S_r.

[Событие $\{T_1 = m_1, T_2 = m_2, \ldots, T_r = m_r\}$ состоит из элементарных событий $\mathbf{x} = (x_1, x_2, \ldots)$, таких, что для $n = m_1 + m_2 + \ldots + m_r + r$

$$[x]_n = (x_1, \ldots, x_n) = (\underbrace{0, \ldots, 0}_{m_1}, 1, \underbrace{0, \ldots, 0}_{m_2}, 1, \ldots, 1, \underbrace{0, \ldots, 0}_{m_r}, 1).$$

Следовательно, при $m_k = 0, 1, 2, \ldots$, $k = 1, 2, \ldots, r$,

$$p_{T_1, T_2, \ldots, T_r}(m_1, m_2, \ldots, m_r) = (1-p)^{m_1} p \cdot (1-p)^{m_2} p \cdot \ldots \cdot (1-p)^{m_r} p.$$

Суммируя по всем m_k, $k = 1, 2, \ldots, r$, а также по всем m_k, кроме одного получаем

$$\sum_{(m_1, m_2, \ldots, m_r)} p_{T_1, T_2, \ldots, T_r}(m_1, m_2, \ldots, m_r) = 1,$$

$$p_{T_k}(m_k) = \sum_{(m_1,\ldots,m_{k-1},m_{k+1},\ldots,m_r)} p_{T_1,\ldots,T_r}(m_1, \ldots, m_r) = (1-p)^{m_k} p,$$

так что все сл. в. T_k имеют одинаковые распределения, или, как говорят, *одинаково распределены*. Отметим, что

$$p_{T_1,T_2,\ldots,T_r}(m_1, m_2, \ldots, m_r) = p_{T_1}(m_1)\, p_{T_2}(m_2) \ldots p_{T_r}(m_r).$$

Событие $\{T_1+\ldots+T_r=m\}$ представим как объединение событий $\{T_1=m_1, \ldots, T_r=m_r\}$ по всевозможным наборам $(m_1, \ldots, m_r)$, таким, что $m_1+\ldots+m_r=m$. Все эти события имеют равные вероятности $(1-p)^m p^r$, а для разных наборов $m_1, \ldots, m_r$ они не пересекаются. Поскольку имеется C_{m+r-1}^m различных решений уравнения $m_1+\ldots+m_r=m$ (см. 1.26), то, воспользовавшись свойством аддитивности вероятностной меры, получаем

$$p_{S_r}(m) = C_{m+r-1}^m (1-p)^m p^r, \quad m = 0, 1, 2, \ldots .$$

Заметим, что

$$\sum_{m=0}^{\infty} p_{S_r}(m) = \sum_{m=0}^{\infty} \sum_{\{m_1+\ldots+m_r=m\}} p_{T_1,\ldots,T_r}(m_1, \ldots, m_r) =$$

$$= \sum_{\{m_1,\ldots,m_r\}} p_{T_1,\ldots,T_r}(m_1, \ldots, m_r) = 1.$$

Распределение вероятностей $p_{S_r}(m)$ называется *отрицательным биномиальным с параметрами* r и p.

Другой способ вычисления распределения вероятностей $p_{S_r}(m)$ состоит в сложении вероятностей $p^r(1-p)^m$ событий $\{\mathbf{x} : [\mathbf{x}]_n = (x_1^0, \ldots, x_n^0)\}$, $n=m+r$, по всем наборам $(x_1^0, \ldots, x_n^0)$, содержащим r единиц, включая единицу на последнем месте (см. 1.24). При этом способе, однако, требуются дополнительные усилия, чтобы установить равенство

$$\sum_{m=0}^{\infty} C_{m+r-1}^m (1-p)^m p^r = 1.]$$

2.18. З а д а ч а. Рассмотрим бесконечную последовательность испытаний Бернулли. Любое $\mathbf{x} = (x_1, x_2, \ldots)$, $x_k=0$ или 1, $k=1, 2, \ldots$, можно записать либо в виде

$$\mathbf{x} = (\underbrace{0, \ldots, 0}_{m_1}, \underbrace{1, \ldots, 1}_{m_2}, \underbrace{0, \ldots, 0}_{m_3}, \underbrace{1, \ldots, 1}_{m_4}, 0, \ldots),$$

либо в виде

$$\mathbf{x} = (\underbrace{1, \ldots, 1}_{m_1}, \underbrace{0, \ldots, 0}_{m_2}, \underbrace{1, \ldots, 1}_{m_3}, \underbrace{0, \ldots, 0}_{m_4}, 1, \ldots)$$

при некоторых целых положительных $m_1, m_2, \ldots$. Введем сл. в. $V_k = V_k(\mathbf{x})$, положив $V_k(\mathbf{x}) = m_k$, $k = 1, 2, \ldots$, и назовем сл. в. V_k длиной k-й серии. Найти совместное и частные распределения вероятностей сл. в. $V_1, V_2, \ldots, V_r$.

[Для упрощения записи возьмем $r=3$. Событие $\{V_1 = m_1, V_2 = m_2, V_3 = m_3\}$ состоит из элементарных событий $\mathbf{x}$, таких, что при $n = m_1 + m_2 + m_3 + 1$ либо

$$[\mathbf{x}_n] = (x_1, \ldots, x_n) = (\underbrace{0, \ldots, 0}_{m_1}, \underbrace{1, \ldots, 1}_{m_2}, \underbrace{0, \ldots, 0}_{m_3}, 1),$$

либо

$$[\mathbf{x}]_n = (\underbrace{1, \ldots, 1}_{m_1}, \underbrace{0, \ldots, 0}_{m_2}, \underbrace{1, \ldots, 1}_{m_3}, 0).$$

Отсюда получаем

$$p_{V_1, V_2, V_3}(m_1, m_2, m_3) = (1-p)^{m_1} p^{m_2} (1-p)^{m_3} p + p^{m_1} (1-p)^{m_2} p^{m_3} (1-p).$$

Суммируя по m_2, m_3, находим

$$p_{V_1}(m_1) = \sum_{m_2=1}^{\infty} \sum_{m_3=1}^{\infty} p_{V_1, V_2, V_3}(m_1, m_2, m_3) = (1-p)^{m_1} p + p^{m_1}(1-p),$$

$$p_{V_2}(m_2) = (1-p)^{m_2-1} p^2 + p^{m_2-1}(1-p)^2, \quad p_{V_3}(m) = p_{V_1}(m).$$

Нетрудно выписать совместное и частные распределения сл. в. $V_1, V_2, \ldots, V_r$ при любом r, действуя аналогично случаю $r=3$. Отметим, что частные распределения сл. в. $V_1, V_3, \ldots$ для серий с нечетными номерами одинаковы, то же верно и для распределений сл. в. $V_2, V_4, \ldots$ с четными номерами, но при $p \neq 1/2$ распределения длин серий с четными и нечетными номерами различны.]

2.19. Пример (задача о разорении игрока). Представим себе следующую игру двух лиц. Проводится бесконечная последовательность испытаний Бернулли. При каждом испытании первый игрок выигрывает условную единицу, если выпадает успех (1), и проигрывает ее при неудаче (0). Пусть a — начальный капитал первого игрока, b — суммарный капитал обоих игроков, $0 < a < b$. Удобно допустить, что игра может продолжаться в долг, так что если $\mathbf{x} = (x_1, x_2, \ldots)$ — элементарное событие схемы испытаний Бернулли, то сл. в. $Y_k = Y_k(\mathbf{x}) = 2x_k - 1$ представляет выигрыш или проигрыш первого игрока в k-й партии, $k = 1, 2, \ldots$, сл. в. $Z_k = a + Y_1 + \ldots + Y_k$ — его капитал (возможно отрицательный) после k-й партии, $k = 1, 2, \ldots$. Последовательность сл. в. $Z_0 = a, Z_1, Z_2 \ldots$ называется *неограниченным простым случайным блужданием*

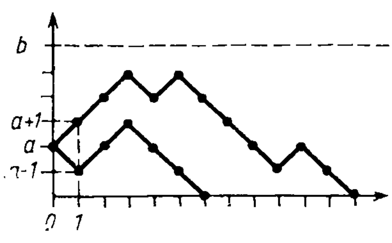

Рис. 12

(ср. 2.11), начинающимся из точки a. Скажем, что первый игрок разорился в момент m, если $Z_m=0$ и $0<Z_k<b$ при $k<m$. Нас будет интересовать вероятность, что первый игрок разорится, или (на языке случайного блуждания) вероятность того, что траектория $(a, z_1, z_2, \ldots)$, $z_k=a+y_1+\ldots+y_k$, $y_j=2x_j-1$, $j, k=1, 2, \ldots$, раньше попадет в точку 0, чем в точку b (рис. 12). Обозначим A_n событие «игрок разорился до n-й партии включительно»:

$$A_n=\{\mathbf{x}: \text{при некотором } m\leqslant n,\ z_m=0,\ 0<z_k<b \text{ для } k<m\}.$$

Множество A_n входит в класс $\mathcal{A}$, на котором определена вероятностная мера $\mathbf{P}$. Множество $A=\bigcup\limits_{n=1}^{\infty} A_n$ состоит из тех и только тех элементарных событий, при которых происходит разорение первого игрока. Однако A не входит в $\mathcal{A}$ и, следовательно, $\mathbf{P}(A)$ еще требует своего определения. Поскольку $A_n\subseteq A_{n+1}$, последовательность $\mathbf{P}(A_n)$ не убывает и потому имеет предел. Кажется вполне естественным положить

$$\mathbf{P}(A)=\lim_{n\to\infty}\mathbf{P}(A_n).$$

Говорят, что неубывающая последовательность множеств $A_1\subseteq\subseteq A_2\subseteq\ldots$ сходится к множеству $A=\bigcup\limits_{n=1}^{\infty} A_n$. Поэтому можно сказать, что $\mathbf{P}(A)$ получено доопределением по непрерывности. Можно получить $\mathbf{P}(A)$ несколько иначе, вводя события

$$B_m=\{\mathbf{x}: z_m=0,\ 0<z_k<b \text{ при } k<m\},\quad m=1, 2, \ldots .$$

События B_m, $m=1, 2, \ldots$, попарно не пересекаются, а их объединение составляет A. Если «применить» свойство счетной аддитивности, то получим

$$\mathbf{P}(A)=\mathbf{P}\Big(\bigcup_{m=1}^{\infty} B_m\Big)=\sum_{m=1}^{\infty}\mathbf{P}(B_m).$$

Так как

$$\mathbf{P}(A_n)=\mathbf{P}\Big(\bigcup_{m=1}^{n} B_m\Big)=\sum_{m=1}^{n}\mathbf{P}(B_m),$$

то приходим к тому же значению $\mathbf{P}(A)$.

2.20*. З а д а ч а. Пусть суммарный капитал b двух игроков фиксирован, обозначим $q_m(a)$ вероятность разорения первого игрока в момент m, если его начальный капитал равнялся a, и $Q(a)=\sum\limits_{m=1}^{\infty} q_m(a)$ — вероятность разорения первого игрока. Вывести рекуррентные соотношения (рис. 12)

$$q_m(a) = pq_{m-1}(a+1) + (1-p)\,q_{m-1}(a-1), \quad m = 1, 2, \ldots,$$

$$Q(a) = pQ(a+1) + (1-p)\,Q(a-1)$$

с граничными условиями $q_k(l) = 0$, $q_k(0) = 1$, $k = 0, 1, 2, \ldots$, $l = 1, 2, \ldots$, $Q(b) = 0$, $Q(0) = 1$. Найти $Q(a)$.

[Пусть $\mathbf{x} = (x_1, x_2, \ldots)$ — элементарное событие схемы бесконечной последовательности испытаний Бернулли,

$$y_k = 2x_k - 1, \; p(x_1, \ldots, x_n) = p^{x_1+\ldots+x_n}(1-p)^{n-x_1-\ldots-x_n}.$$

Введем обозначение

$$B_a^{(m)} = \{(x_1, \ldots, x_m) : a + y_1 + \ldots + y_m = 0, \quad 0 < a + y_1 + \ldots$$
$$\ldots + y_k < b \text{ при } k < b\}.$$

По определению вероятностной меры $\mathbf{P}$ на классе множеств $\mathscr{A}$ (см. 2.15)

$$q_m(a) = \sum_{(x_1, \ldots, x_m) \in B_a^{(m)}} p(x_1, \ldots, x_m) =$$

$$= \sum_{x_1 \in \{0,1\}} p(x_1) \sum_{(x_1, \ldots, x_m) \in B_a^{(m)}} p(x_2, \ldots, x_m) =$$

$$= (1-p) \sum_{(0, x_2, \ldots, x_m) \in B_a^{(m)}} p(x_2, \ldots, x_m) + p \sum_{(1, x_2, \ldots, x_m) \in B_a^{(m)}} p(x_2, \ldots x_m).$$

Но $\{(x_2, \ldots, x_m) : (0, x_2, \ldots, x_m) \in B_a^{(m)}\} = \{(x_2, \ldots, x_m) : a - 1 + y_2 + \ldots + y_m = 0, \; 0 < a - 1 + y_2 + \ldots + y_k < b$ при $k < b\} = B_{a-1}^{(m-1)}$ и аналогично $\{(x_2, \ldots, x_m) : (1, x_2, \ldots, x_m) \in B_a^{(m)}\} = B_{a+1}^{(m-1)}$. Таким образом, при $m > 1$, $1 < a < b-1$, получаем уравнение

$$q_m(a) = (1-p)\,q_{m-1}(a-1) + pq_{m-1}(a+1).$$

Аналогичные рассуждения приводят к соотношениям

$$q_m(1) = pq_{m-1}(2), \; m > 1, \quad q_m(b-1) = (1-p)\,q_{m-1}(b-2),$$

и, кроме того, $q_1(1) = 1-p$, $q_1(a) = 0$ при $1 < a < b$. Все эти соотношения сводятся к предыдущему уравнению, рассматриваемому в области $m \geqslant 1$, $1 \leqslant a \leqslant b-1$, с граничными условиями $q_k(0) = 1$, $q_k(l) = 0$, $k = 0, 1, 2, \ldots$, $l = 1, 2, \ldots$.

Суммируя по m, находим

$$\sum_{m=1}^{\infty} q_m(a) = (1-p) \sum_{m=1}^{\infty} q_{m-1}(a-1) + p \sum_{m=1}^{\infty} q_{m-1}(a+1),$$

или

$$Q(a) = (1-p)\,Q(a-1) + pQ(a+1)$$

с граничными условиями $Q(0)=1$, $Q(b)=0$. Перепишем полученное уравнение в виде

$$p(Q(j)-Q(j+1))=(1-p)(Q(j-1)-Q(j)),$$

или

$$Q(j)-Q(j+1)=\lambda(Q(j-1)-Q(j)), \quad \lambda=(1-p)/p.$$

Итерацией по j с учетом $Q(0)=1$ имеем

$$Q(j)-Q(j+1)=\lambda^j(1-Q(1)).$$

Суммируя по j от 0 до $k-1$, получаем

$$1-Q(k)=(1+\lambda+\ldots+\lambda^{k-1})(1-Q(1)).$$

Полагая $k=b$ и учитывая, что $Q(b)=0$, имеем

$$1=(1+\lambda+\ldots+\lambda^{b-1})(1-Q(1)).$$

Подставляя $1-Q(1)$ из последнего равенства в предыдущее, находим

$$Q(k)=(\lambda^k+\lambda^{k-1}+\ldots+\lambda^{b-1})/(1+\lambda+\ldots+\lambda^{b-1}), \quad \lambda=(1-p)/p.]$$

§ 3.

НЕПРЕРЫВНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ

В экспериментальных науках информацию обычно представляют десятичными дробями с некоторым числом знаков после запятой, обусловленным точностью измерений. Казалось бы, дискретная вероятностная модель должна адекватно описывать любой реальный опыт вероятностного характера. Однако модели с непрерывным множеством состояний в теории вероятностей, как и в математике вообще, имеют несомненное преимущество перед дискретными ввиду их более ясной логической структуры и значительно более развитого аппарата их исследования.

В качестве модельного примера рассмотрим опыт по выбору наудачу точки из отрезка [0, 1]. Пространство элементарных событий состоит здесь из несчетного множества бесконечных десятичных дробей $0,x_1x_2\ldots$, которые следует считать равновозможными. Легко понять, что единственный способ удовлетворить такому требованию — положить вероятности всех элементарных событий равными нулю, что не приводит к содержательной вероятностной модели опыта. Представление о равновозможности исходов опыта наводит, с другой стороны, на мысль, что все 10^n десятичных приближений $0,x_1x_2\ldots x_n$ результатов опыта надо считать равновероятными. Иначе говоря, событие {случайно выбранная точка из отрезка [0, 1] попала в промежуток $[0,x_1x_2\ldots x_n, 0,x_1x_2\ldots x_n+10^{-n})$} должно иметь вероятность 10^{-n}, равную длине промежутка. Принимая свойство аддитивности вероятности, получаем, что вероятность попадания случайно выбранной точки в любой промежуток с концами 0, $x_1 \ldots x_n<0$, $y_1 \ldots y_n$ следует положить равной его длине. Наконец, вероятность попадания в любой интервал $(x, y)\subseteq[0, 1]$ можно положить равной пределу вероятностей попадания в интервалы $(0, x_1 \ldots x_n$, $0, y_1 \ldots y_n)$, т. е. равной длине $y—x$ этого интервала. Итак, мы приходим к вероятностной модели, в которой пространство элементарных событий есть множество действительных чисел от 0 до 1, а вероятностная мера определена на интервалах и совпадает с длиной. Те же самые соображения ведут от дискретной модели повторного случайного выбора к модели *повторного случайного выбора точки из отрезка* [0, 1]. В качестве пространства элементарных событий здесь выступает n-мерный куб $[0, 1]^n$, а вероятностная мера любого параллелепипеда $(a_1, b_1)\times(a_2, b_2)\times\times\ldots\times(a_n, b_n)\subseteq[0, 1]^n$ равняется его объему: $(b_1—a_1)\times\times\ldots\times(b_n—a_n)$.

Класс событий, состоящий из интервалов при однократном выборе и параллелепипедов при повторном, недостаточен для описания различных свойств результата опыта. Расширение области определения вероятностной меры в рассматриваемом случае сводится к определению понятия длины, площади, объема, n-мерного объема для возможно более широкого класса множеств. Откладывая детальное изложение этих вопросов до последующих разделов книги, будем пользоваться пока теми соображениями, которые сообщаются в курсах математического анализа. Отметим, в частности, что точка имеет нулевую длину, отрезок — нулевую площадь и т. д., поэтому любой промежуток $I_{a,b}=(a, b)$, $[a, b)$, $(a, b]$, $[a, b]\subseteq[0, 1]$ имеет вероятность, равную его длине, любой прямоугольник $I_{a_1,b_1}\times I_{a_2,b_2}$ в опыте с двукратным случайным выбором имеет вероятность, равную его площади, и т. д.

3.1. Задача. Две точки брошены наудачу на отрезок [0, 1]. Какова вероятность, что из трех отрезков, на которые они разбивают отрезок [0, 1], можно составить треугольник.
[Пространство элементарных событий состоит из точек (x_1, x_2), принадлежащих квадрату $[0, 1]\times[0, 1]$. Обозначим $x_{(1)}=\min\{x_1, x_2\}$, $x_{(2)}=\max\{x_1, x_2\}$ и запишем условие образования треугольника из отрезков длины $x_{(1)}$, $x_{(2)}-x_{(1)}$, $1-x_{(2)}$:

$$x_{(1)}+(x_{(2)}-x_{(1)})>1-x_{(2)}, \quad x_{(1)}+(1-x_{(2)})>x_{(2)}-x_{(1)},$$
$$(x_{(2)}-x_{(1)})+(1-x_{(2)})>x_{(1)},$$

или

$$x_{(2)}>1/2, \quad x_{(2)}-x_{(1)}<1/2, \quad x_{(1)}<1/2.$$

При $x_{(1)}=x_1$, $x_{(2)}=x_2$ область $\{(x_1, x_2): x_2>1/2,\ x_2-x_1<1/2,\ x_1<1/2\}$ представляет собой треугольник с площадью 1/8, при $x_{(1)}=x_2$, $x_{(2)}=x_1$, получаем симметричный треугольник (рис. 13), так что искомая вероятность равна 1/4.]

3.2. Задача. На отрезок [0, 1] брошены наудачу n точек.

Какова вероятность, что: (I) крайняя правая точка окажется левее заданного $t \in [0, 1]$; (II) крайняя левая точка окажется левее некоторого $s \in [0, 1]$; (III) все точки поместятся на заданный отрезок $[s, t] \subseteq [0, 1]$.

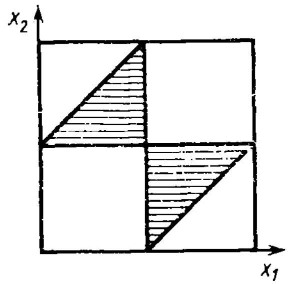

Рис. 13

[Событие (I) означает, что точка $(x_1, \ldots, x_n)$ попала в куб $[0, t]^n$ и имеет вероятность t^n. Событие (II) является дополнительным (или *противоположным*) событию {все n точек оказались правее s}, которое имеет вероятность $(1-s)^n$. Отсюда вероятность события (II) равна $1-(1-s)^n$. Событие (III) ет вероятность $(t-s)^n$.]

3.3. Задача. На отрезок [0, 1] брошено случайно n точек. Найти вероятность того, что слева от заданного $t \in [0, 1]$ попадет ровно k точек, $k=0, 1, \ldots, n$.

[В кубе $[0, 1]^n$ выделяется C_n^k непересекающихся областей вида $\{(x_1, \ldots, x_n) : x_{i_1} < t, \ldots, x_{i_k} < t$, остальные $x_j > t\}$, $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$, каждая из которых имеет вероятность $t^k(1-t)^{n-k}$, откуда искомая вероятность равна $C_n^k t^k (1-t)^{n-k}$.]

Вероятностная модель опыта по повторному выбору точек из отрезка [0, 1] состоит из пространства элементарных событий $\Omega = [0, 1]^n (n=1, 2, \ldots)$ и вероятностной меры $\mathbf{P}(A)$, заданной на системе $\mathcal{A}$ подмножеств $A \subseteq [0, 1]^n$, для которых определено понятие n-мерного объема. Понятия длины, площади, объема, n-мерного объема не могут быть распространены на любые подмножества с сохранением характерных свойств этих понятий, и потому класс множеств $\mathcal{A}$ не включает все подмножества пространства элементарных событий Ω, как это имеет место в дискретной модели. *Событиями* называются лишь те подмножества Ω, которые входят в $\mathcal{A}$ и для которых, следовательно, определена вероятность. Класс множеств $\mathcal{A}$ включает всевозможные параллелепипеды, для которых

$$\mathbf{P}(I_{a_1,b_1} \times \ldots \times I_{a_n,b_n}) = (b_1-a_1) \cdot \ldots \cdot (b_n-a_n),$$

$$I_{a_i,b_i} \subseteq [0, 1], \quad i=1, \ldots, n.$$

Предполагаем, что класс $\mathcal{A}$ — алгебра множеств, т. е. замкнут относительно перехода к дополнению, объединению и пересечению (см. 2.16), а вероятностная мера $\mathbf{P}$ на $\mathcal{A}$ аддитивна

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B), \quad AB = \varnothing, \ A, B \in \mathcal{A}.$$

Отметим, что из аддитивности меры $\mathbf{P}$ вытекают часто используемые свойства: $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$, $\mathbf{P}(B) \leqslant \mathbf{P}(A)$ при $B \subseteq A$ и др. (см. 2.2).

В согласии со сказанным *вероятностным пространством* в рассматриваемом случае называют совокупность из трех объектов: множества Ω, класса $\mathscr{A}$ подмножеств Ω, называемых событиями, и вероятностной меры $\mathbf{P}$, заданной на $\mathscr{A}$. Действительная функция $X(\omega)$ называется *случайной величиной* (сл. в.), если для любого промежутка $I_{a,b}$, $-\infty \leqslant a < b \leqslant \infty$, подмножество $\{\omega : X(\omega) \in I_{a,b}\}$ входит в $\mathscr{A}$. Понятно, что это требование является необходимым для того, чтобы функция $X(\omega)$ могла служить объектом изучения теории вероятностей. Случайная величина (сл. в.) $X(\omega)$ порождает вероятностную меру

$$\mathbf{P}_X(I_{a,b}) = \mathbf{P}(\omega : X(\omega) \in I_{a,b})$$

на классе всех промежутков, которую называют распределением вероятностей случайной величины X. Набор $\mathbf{X}(\omega) = (X_1(\omega), \ldots$ $\ldots, X_n(\omega))$ случайных величин порождает вероятностную меру на n-мерных параллелепипедах

$$\mathbf{P}_{\mathbf{X}}(I_{a_1,b_1} \times \ldots \times I_{a_n,b_n}) = \mathbf{P}(\omega : X_1(\omega) \in I_{a_1,b_1}, \ldots, X_n(\omega) \in I_{a_n,b_n}),$$

которую называют *совместным распределением вероятностей случайных величин* $X_1, \ldots, X_n$.

Функция $F_X(x) = \mathbf{P}_X((-\infty, x])$ действительного переменного x называется *функцией распределения* (ф. р.) сл. в. X. Легко видеть, что

$$\mathbf{P}_X((a, b]) = F_X(b) - F_X(a)$$

для любых $a < b$. Функция n действительных переменных

$$F_X(x_1, \ldots, x_n) = \mathbf{P}_X((-\infty, x_1] \times \ldots \times (-\infty, x_n])$$

называется *совместной функцией распределения сл. в.* $X_1, \ldots, X_n$. В модели n-кратного случайного выбора точек из отрезка $[0, 1]$ сл. в. $X_i(\omega) = x_i$, $\omega = (x_1, \ldots, x_n)$, имеет распределение

$$\mathbf{P}_{X_i}(I_{a,b}) = b - a, \quad I_{a,b} \subseteq [0, 1],$$

которое называется *равномерным распределением* на отрезке $[0, 1]$. Функция распределения X_i равна 0 при $x < 0$, равна 1 при $x > 1$ и равна x при $0 \leqslant x \leqslant 1$. Совместная функция распределения сл. в. $X_1, \ldots, X_n$ равна

$$F_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n).$$

Из 3.2 вытекает, что сл. в. $U(\omega) = \max\{x_1, \ldots, x_n\}$, $\omega = (x_1, \ldots, x_n)$, имеет ф. р. $F_U(u) = u^n$ при $0 \leqslant u \leqslant 1$, сл. в. $V(\omega) = \min(x_1, \ldots, x_n)$ имеет ф. р. $F_V(v) = 1 - (1 - v)^n$ при $0 \leqslant v \leqslant 1$, $F_U(t) = F_V(t) = 0$ при $t < 0$, $F_U(t) = F_V(t) = 1$ при $t > 1$. Сл. в. $Z_t(\omega)$ — число точек, попавших левее заданного $t \in (0, 1)$, имеет биномиальное распределение вероятностей (см. 3.3.):

$$\mathbf{P}(\omega : Z_t(\omega) = k) = C_n^k t^k (1 - t)^{n-k}, \quad k = 0, 1, \ldots, n.$$

Ее функция распределения

$$\mathbf{P}(\omega : Z_t(\omega) \leqslant z) = \sum_{k=0}^{[z]} C_n^k t^k (1-t)^{n-k}, \quad 0 \leqslant z \leqslant n,$$

является кусочно-постоянной неубывающей функцией со скачками в точках $k=0, 1, \ldots, n$.

Случайная величина и ее распределение вероятностей называются *непрерывными*, если непрерывна функция распределения.

3.4. Задача. На отрезок $[0, 1]$ наудачу брошено n точек, $X_{(k)}(\omega)$ — положение k-й по порядку слева направо точки ($X_{(1)}(\omega)$ — крайняя левая, $X_{(n)}(\omega)$ — крайняя правая точки). Найти функцию распределения сл. в. $X_{(k)}(\omega)$.

[Событие $\{\omega : X_{(k)}(\omega) \leqslant x\}$ происходит тогда и только тогда, когда число $Z_x(\omega)$ точек, попавших слева от x, больше либо равно k. Отсюда и из 3.3 получаем

$$F_{X_{(k)}}(x) = \mathbf{P}(\omega : Z_x(\omega) \geqslant k) = \sum_{j=k}^{n} C_n^j x^j (1-x)^{n-j}, \quad 0 \leqslant x \leqslant 1.]$$

3.5. Задача. В обозначениях задачи 3.4 найти совместную функцию распределения сл. в. $X_{(1)}, X_{(n)}$.

[При $0 \leqslant y < x \leqslant 1$ имеем равенство

$$\{\omega : X_{(1)}(\omega) \leqslant x, \ X_{(n)}(\omega) \leqslant y\} = \{\omega : X_{(n)}(\omega) \leqslant y\},$$

откуда получаем $F_{X_{(1)}, X_{(n)}}(x, y) = y^n$ при $0 \leqslant y < x \leqslant 1$. При $0 \leqslant x < y \leqslant 1$ воспользуемся соотношением

$$\{\omega : X_{(1)}(\omega) \leqslant x, \ X_{(n)}(\omega) \leqslant y\} =$$
$$= \{\omega : X_{(n)}(\omega) \leqslant y\} \ominus \{\omega : X_{(1)}(\omega) > x, \ X_n(\omega) \leqslant y\}.$$

Поскольку

$$\{\omega : X_{(1)}(\omega) > x, \ X_{(n)}(\omega) \leqslant y\} = \{\omega : x < X_i(\omega) \leqslant y, \ i = 1, \ldots, n\},$$

находим отсюда

$$F_{X_{(1)}, X_{(n)}}(x, y) = F_{X_{(n)}}(y) - (F_{X_1}(y) - F_{X_1}(x))^n = y^n - (y-x)^n.]$$

3.6. Задача. Показать, что

$$\mathbf{P}((a_1, b_1] \times (a_2, b_2]) = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2).$$

$$[\mathbf{P}((a_1, b_1] \times (a_2 \ b_2]) = \mathbf{P}((-\infty, b_1] \times (a_2, b_2]) -$$
$$-\mathbf{P}((-\infty, a_1] \times (a_2, b_2]) = \mathbf{P}((-\infty, b_1] \times (-\infty, b_2]) -$$
$$-\mathbf{P}((-\infty, b_1] \times (-\infty, a_2]) - (\mathbf{P}((-\infty, a_1] \times (-\infty, b_2]) -$$
$$-\mathbf{P}((-\infty, a_1] \times (-\infty, a_2])) = F(b_1, b_2) - F(b_1, a_2) -$$
$$-F(a_1, b_2) + F(a_1, a_2).]$$

Распределению вероятностей полезно сопоставить физический образ — распределение единичной массы. Равномерному дискретному распределению на множестве чисел k/N, $k=0, 1, \ldots, N-1$, соответствуют точечные массы N^{-1}, помещенные в точки k/N отрезка $[0, 1]$. Равномерное непрерывное распределение единичной массы на отрезке $[0, 1]$ не имеет точек сосредоточения массы, а любой отрезок $[\alpha, \beta] \subseteq [0, 1]$ имеет массу, равную его длине. Его можно представить себе как предельный случай равномерного дискретного распределения при $N \to \infty$.

Неравномерное распределение единичной массы на некотором отрезке $[a, b]$ можно получить, производя деформацию стержня $[0, 1]$ с единичной равномерно распределенной массой. Пусть $y=F(x)$, $a \leqslant x \leqslant b$, — возрастающая непрерывная функция, $F(a)=0$, $F(b)=1$ (рис. 14), $x=F^{-1}(y)$ — функция, обратная к ней.

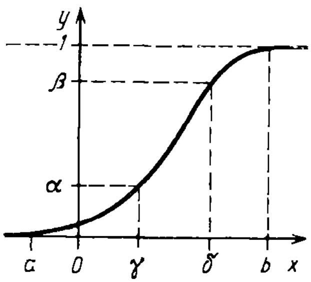

Рис. 14

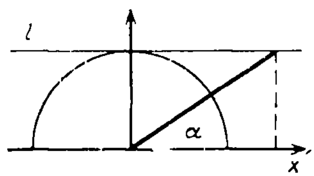

Рис. 15

Отрезок $[\alpha, \beta] \subseteq [0, 1]$ при деформации $y \to F^{-1}(y)$ отрезка $[0, 1]$ переходит в отрезок $[\gamma, \delta]$, $\gamma=F^{-1}(\alpha)$, $\delta=F^{-1}(\beta)$ с сохранением массы $\beta-\alpha$. Иначе говоря, произвольный отрезок $[\gamma, \delta] \subseteq [a, b]$ имеет массу $F(\delta)-F(\gamma)$. Средняя плотность $(F(\delta)-F(\gamma))/(\delta-\gamma)$ распределения массы на отрезке $[\gamma, \delta]$ при стягивании отрезка $[\gamma, \delta]$ к точке x стремится к производной $F'(x)=f(x)$, если она существует. Ее называют плотностью распределения массы в точке x. Заметим, что если производная существует везде, то

$$F(\gamma)-F(\delta)=\int_{\gamma}^{\delta} f(x)\,dx.$$

Равномерное распределение массы на отрезке $[a, b]$ задается функцией $F(x)=(x-a)/(b-a)$ с плотностью $f(x)=1/(b-a)$, $x \in [a, b]$. Взяв произвольную неотрицательную интегрируемую функцию $f(x)$, $-\infty < x < \infty$, такую, что

$$\int_{-\infty}^{\infty} f(x)\,dx = 1,$$

можно задать распределение единичной массы на прямой с помощью *функции распределения* (ф. р.)

$$F(x)=\int_{-\infty}^{x} f(y)\,dy,\quad -\infty<x<\infty.$$

В любой точке непрерывности функции $f(x)$ имеем $F'(x)=f(x)$, и потому функция $f(x)$, $-\infty<x<\infty$, называется *плотностью распределения* массы.

Переходя от физической модели к вероятностной, необходимо лишь заменить в предыдущих рассуждениях массу физическую на «массу вероятностную», или вероятность. С формальной точки зрения в обоих случаях речь идет об одном и том же математическом понятии — числовой мере на прямой, различие состоит только в интерпретации, но оно оказывается решающим при определении круга задач и методов их исследования. Распределение вероятностей на прямой, задаваемое плотностью, называется *абсолютно непрерывным*. Для абсолютно непрерывных распределений вопрос о расширении области определения вероятностной меры с промежутков на более сложные множества можно отнести к теории интеграла, полагая

$$\mathbf{P}(A)=\int_{A} f(x)\,dx$$

на классе $\mathcal{A}$ подмножеств $A\subseteq R$, для которых интеграл является аддитивной функцией множества.

3.7. Задача. Из начала координат проведен луч в случайном направлении α, $0<\alpha<\pi$, до пересечения с прямой l, параллельной оси абсцисс и отстоящей от нее на единичном расстоянии (рис. 15). Предполагая, что угол α равномерно распределен на отрезке $[0,\ \pi]$, найти функцию и плотность распределения вероятностей абсциссы точки пересечения луча и прямой l.

[Отображение $x=\operatorname{ctg}\alpha$ переводит промежуток $[\alpha,\ \pi)$ в $(-\infty,\ x]$, откуда функция распределения равна $F(x)=\pi^{-1}(\pi-\alpha)=\pi^{-1}(\pi-\operatorname{arcctg} x)$, $-\infty<x<\infty$. Дифференцируя, находим плотность $f(x)=F'(x)=(\pi(1+x^2))^{-1}$. Распределение с плотностью $f(x)$ носит имя Коши.]

3.8. Задача. Пусть сл. в. $X(\omega)$ равномерно распределена на отрезке $[0,\ 1]$. Найти функции и плотности распределений сл. в.

$$X^2,\quad \sqrt{X},\quad \sin^2(\pi X/2).$$

$$[F_1(x)=\mathbf{P}(\omega: X^2(\omega)\leqslant x)=\mathbf{P}\left(\omega: X(\omega)\leqslant \sqrt{x}\right)=\sqrt{x},\ 0\leqslant x\leqslant 1,$$

$$f_1(x)=F_1'(x)=1/(2\sqrt{x});\ F_2(x)=\mathbf{P}\left(\omega:\sqrt{X(\omega)}\leqslant x\right)=x^2,\ 0\leqslant x\leqslant 1,$$

$$f_2(x)=2x;\ F_3(x)=\mathbf{P}(\omega:\sin^2(\pi X(\omega)/2)\leqslant x)=2\pi^{-1}\arcsin\sqrt{x},$$

$$0\leqslant x\leqslant 1,\ f_3(x)=\pi^{-1}(x(1-x))^{-1/2},\ 0<x<1.]$$

3.9. Задача. Пусть $\omega=0$, $x_1 x_2 \ldots$ есть двоичная запись числа $\omega\in[0, 1]$. В опыте по случайному выбору точки из отрезка $[0, 1]$ введем бесконечную последовательность сл. в. $X_k(\omega)=x_k$, $k=1, 2, \ldots$. Показать, что последовательность сл. в. $X_1, X_2, \ldots$ представляет собой испытания Бернулли с вероятностью успеха 1/2.

$$[\mathbf{P}(\omega : X_k(\omega)=x_k, \quad k=1, \ldots, n)=$$
$$=\mathbf{P}([0, x_1 \ldots x_n, 0, x_1 \ldots x_n+2^{-n}))=2^{-n}.]$$

3.10. Задача. Найти плотность распределения вероятностей сл. в. $X_{(k)}$ в 3.4.

$$\Bigg[f(x)=\frac{d}{dx}\sum_{i=k}^{n} C_n^i x^i (1-x)^{n-i}=\sum_{i=k}^{n}\frac{n!}{i!\,(n-i)!}\, i x^{i-1}(1-x)^{n-i}-$$
$$-\sum_{i=k}^{n}\frac{n!}{i!\,(n-i)!}(n-i)\,x^i(1-x)^{n-i-1}=\sum_{i=k}^{n} nC_{n-1}^{i-1}x^{i-1}(1-x)^{(n-1)-(i-1)}-$$
$$-\sum_{i=k}^{n-1} nC_{n-1}^{i}x^i(1-x)^{n-1-i}=nC_{n-1}^{k-1}x^{k-1}(1-x)^{n-k}.\Bigg]$$

3.11. Пример (пуассоновский процесс). Представим себе, что на временной оси регистрируются случайные моменты появления некоторых событий. Фиксируя каким-либо образом точку отсчета 0, примем для этих моментов обозначение (рис. 16)

$$\ldots<\tau_{-2}<\tau_{-1}<0<\tau_1<\tau_2<\ldots<\tau_n<\tau_{n+1}<\ldots.$$

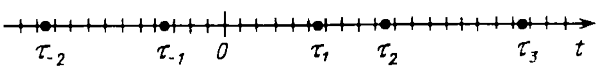

Рис. 16

Вероятностную природу рассматриваемого случайного потока событий определим путем ее приближенного описания моделью испытаний Бернулли. Именно предположим, что временная ось разбита на промежутки малой длины h, на каждом из которых производится испытание Бернулли с некоторой малой вероятностью успеха p_h, так что в целом имеется бесконечная в обе стороны последовательность испытаний Бернулли. Успех в испытании, проведенном на интервале $(kh, (k+1)h)$, $k=0, \pm1, \pm2, \ldots$, будем трактовать как появление события на этом интервале. Получаемый таким путем *поток событий* примем за приближение исходного потока, считая, что при $h\to0$, $p_h\to0$ это приближение становится все более точным.

Для испытаний Бернулли с положительными номерами обозначим $T_i^{(h)}$, $i=2, 3, \ldots$ число неудач между $(i-1)$-м и i-м успехами и положим $T_1^{(h)}$ равным числу неудач до первого успеха. Аналогично определяем сл. в. $T_i^{(h)}$, $i=-1, -2, \ldots$, по испытаниям Бернулли с отрицательными номерами. Функция распределения сл. в. $hT_1^{(h)}$ равна (см. 2.17)

$$F^{(h)}(x) = \sum_{i=0}^{[x/h]} (1-p_h)^i p_h = 1-(1-p_h)^{[x/h]+1},$$

где $[a]$ обозначает целую часть числа a. Вероятность $F^{(h)}(x)$ того, что $hT_1^{(h)} \in [0, x]$ по замыслу должна при $h \to 0$ стремиться к вероятности $F(x)$ того, что $\tau_1 \in [0, x]$. Предполагая, что при некотором $x>0$ ф. р. $F(x)$ удовлетворяет неравенству $0<F(x)<1$, из соотношения

$$1-(1-p_h)^{[x/h]+1} \to F(x), \quad h \to 0,$$

выводим, что обязательно $p_h \to 0$ при $h \to 0$ и, более того, $p_h/h \to \lambda$, где $\lambda>0$ — некоторое число. В самом деле, переходя к логарифмам, имеем

$$([x/h]+1)\ln(1-p_h) \sim -(x/h)p_h \to \ln(1-F(x)).$$

Итак, наводящие соображения о структуре случайного потока событий $\ldots<\tau_{-2}<\tau_{-1}<\tau_1<\tau_2<\ldots$ приводят нас к условию

$$F_{\tau_1}(x) \equiv F(x) = 1-e^{-\lambda x}, \quad 0 \leqslant x < \infty.$$

Распределение вероятностей с ф. р. $F(x)$ и плотностью $f(x)=\lambda e^{-\lambda x}$ при $x>0$, $f(x)=0$ при $x \leqslant 0$ называется *экспоненциальным* с параметром λ.

Обозначим $N_{s,t}$ число случайных событий, наблюдавшихся на временном промежутке $(s, t]$. Приближением к $N_{0,t}$ является сл. в. $N_{0,t}^{(h)}$ — число успехов в испытаниях Бернулли с номерами от 1 до $[t/h]$, которая имеет биномиальное распределение вероятностей с параметрами $n=[t/h]$, $p=p_h$. Переходя к пределу при $h \to 0$, $p_h/h \to \lambda$, фиксированном k, имеем

$$C_n^k p^k (1-p)^{n-k} = (1/k!)\, n(n-1)\ldots(n-k+1)\, p^k (1-p)^{n-k} =$$
$$= \frac{(np)^k}{k!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)(1-p)^{-k}(1-p)^n \to \frac{(\lambda t)^k}{k!}e^{-\lambda t},$$

поскольку

$$np = [t/h]\,p_h \sim tp_h/h \to \lambda t, \quad (1-p_h)^{-k} \to 1,$$
$$1 > (1-1/n)\ldots(1-(k-1)/n) \geqslant (1-k/n)^k \to 1,$$
$$(1-p_h)^{[t/h]} = (1-\lambda h+o(h))^{(t/h)(1+o(1))} \to e^{-\lambda t}.$$

Итак, сл. в. $N_{0,t}$ должна по замыслу иметь распределение вероятностей

$$p(k;\ \lambda t) = ((\lambda t)^k/k!)e^{-\lambda t},\quad k=0,\ 1,\ 2,\ \dots .$$

Распределение $p(k;\ \lambda)$ называется *пуассоновским* с параметром λ. Утверждение

$$C_n^k p^k (1-p)^{n-k} \to p(k;\ \lambda)$$

при $n\to\infty$, $p\to 0$, $np\to\lambda$ и любом фиксированном $k=0,\ 1,\ 2,\ \dots$ называют *теоремой Пуассона.*

Пуассоновский поток случайных событий $\dots,\ \tau_{-2},\ \tau_{-1},\ \tau_1,\ \tau_2,\ \dots$ пока не получил у нас точного вероятностного описания. Однако приведенная конструкция позволяет легко определить совместное распределение сл. в. $T_1=\tau_1$, $T_2=\tau_2-\tau_1$, $\dots$, $T_n=\tau_n-\tau_{n-1}$, $T_{-1}=$ $=\tau_1$, $T_{-2}=\tau_{-1}-\tau_{-2}$, $\dots$, $T_{-n}=\tau_{-n}-\tau_{-n-1}$ при любом n, что дает ключ к построению вероятностного пространства, на котором могут быть определены такие случайные величины.

3.12. Задача. В обозначениях примера 3.11 показать, что совместная функция распределения сл. в. $hT_{-n}^{(h)}, \dots, hT_{-1}^{(h)}, hT_1^{(h)}$, $\dots$, $hT_n^{(h)}$ имеет своим пределом функцию $F(t_{-n})\dots F(t_{-1})F(t_1)\dots$ $\dots F(t_n)$, где $F(x)=1-e^{-\lambda x}$, $0\leqslant x<\infty$, $F(x)=0$, $x<0$.

[Возьмем для краткости $n=1$ и, используя 2.17, запишем совместную функцию распределения сл. в. $hT_{-1}^{(h)}$, $hT_1^{(h)}$:

$$F^{(h)}(x,\ y) = \sum_{i=0}^{[x/h]}\sum_{j=0}^{[y/h]} (1-p_h)^i p_h (1-p_h)^j p_h,$$

которая ввиду 3.11 сходится при $h\to 0$, $p_h/h\to\lambda$ к $F(x)F(y)$.]

3.13. Задача. В обозначениях примера 3.11 найти предельную функцию распределения сл. в. $hT_{-1}^{(h)}+hT_1^{(h)}$ и ее плотность.

[Сл. в. $S_h=T_{-1}^{(h)}+T_1^{(h)}$ имеет отрицательное биномиальное распределение (см. 2.17): $p_{S_h}(k)=(k+1)(1-p_h)^k p_h^2$, $k=0,\ 1,\ 2,\ \dots$. Ф.р. F_{S_h} сл. в. S_h равна

$$F_{S_h}(x)=\sum_{k=0}^{[x]} p_{S_h}(k) = p_h^2\sum_{k=0}^{[x]}(k+1)(1-p_h)^k.$$

Заметив, что

$$\sum_{k=0}^{n}(k+1)t^k = \frac{d}{dt}\sum_{k=0}^{n} t^{k+1} = \frac{d}{dt}\left(\frac{1-t^{n+2}}{1-t}-1\right)=$$

$$=(1-t)^{-2}(1-t^{n+2}-(n+2)(1-t)t^{n+1}),$$

получаем при $x>0$, $n=[x/h]$

$$F_{hS_h}(x)=F_{S_h}(x/h)=1-(1-p_h)^{n+2}-(n+2)p_h(1-p_h)^{n+1}\to$$

$$\to 1-e^{-\lambda x}(1+\lambda x),\quad f_{hS_h}(x)\to\lambda^2 x e^{-\lambda x},\quad h\to 0.\ ]$$

Заслуживает внимания то обстоятельство, что интервал между соседними событиями, накрывающий точку 0, имеет распределение вероятностей, отличное от экспоненциального, в то время как все остальные интервалы (τ_{k-1}, τ_k), $k=-1, \pm 2, \ldots$, имеют одно и то же экспоненциальное распределение. Это так называемый *парадокс времени ожидания*: начав наблюдение за пуассоновским потоком событий в некоторый момент, обозначенный у нас нулем, мы вправе ожидать ввиду однородности потока по времени, что время до первого появления события должно быть меньше чем «обычное» расстояние между соседними событиями, так как часть этого времени уже протекла до момента начала наблюдения. Однако, как мы видим, это не так.

3.14. Задача. Рассмотрим отрицательное биномиальное распределение

$$p^{(h)}(kh)=C^k_{r+k-1}p_h^r(1-p_h)^k, \quad k=0,\ 1,\ 2,\ \ldots,$$

сосредоточенное в точках прямой вида kh, $k=0, 1, 2, \ldots$, параметр p_h которого зависит от h так, что $p_h/h\to\lambda>0$ при $h\to 0$. Положим

$$\mathbf{P}^{(h)}((a,\ b])=\sum_{k=[a/h]+1}^{[b/h]} p^{(h)}(kh), \quad 0\leqslant a<b,$$

равной суммарной вероятностной массе, лежащей на полуинтервале $(a, b]$. Показать, что при $h\to 0$ имеют место следующие соотношения:

$$p^{(h)}(kh)\sim(k^{r-1}/(r-1)!)\,\lambda^r h^r e^{-\lambda kh}$$

равномерно по $k\to\infty$, $kh^2\to 0$ (т. е. равномерно по $k>\psi(h)$, где $\psi(h)\to\infty$, $\psi(h)\,h^2\to 0$),

$$\mathbf{P}^{(h)}((a,\ b])\to\int\limits_a^b f(x)\,dx,\ f(x)=(\lambda^r/(r-1)!)\,x^{r-1}e^{-\lambda x}$$

для любых $0<a<b$.

[$p^{(h)}(kh)=(k+r-1)(k+r-2)\ldots(k+1)/(r-1)!\ p_h^r(1-p_h)^k\sim hf(kh)$.

$$\mathbf{P}^{(h)}((a,\ b])\sim\sum_{k=[a/h]+1}^{[b/h]} hf(kh)\to\int\limits_a^b f(x)\,dx.]$$

Функция

$$\Gamma(\alpha)=\int\limits_0^\infty x^{\alpha-1}e^{-x}\,dx$$

действительного переменного $\alpha>0$ называется *гамма-функцией* Эйлера. Интегрированием по частям легко получить соотношение

$\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$, откуда, в частности, получаем $\Gamma(n+1)=n!$. Отметим еще известную из анализа формулу $\Gamma(1/2)=\sqrt{\pi}$. Функция

$$g(x)=(1/\Gamma(\alpha))x^{\alpha-1}e^{-x},\quad x>0,\quad g(x)=0,\quad x\leqslant 0,$$

является плотностью распределения вероятностей и называется *гамма-плотностью* с параметром формы $\alpha>0$; экспоненциальное распределение является частным случаем гамма-распределения при $\alpha=1$. На рис. 17 представлены графики $g(x)$ для нескольких значений α. В задаче 3.14 плотность предельного распределения равна $f(x)=\lambda g(\lambda x)$, где $g(x)$ — гамма-плотность с параметром формы $\alpha=r$. Параметр λ плотности $\lambda g(\lambda x)$ называется масштабным.

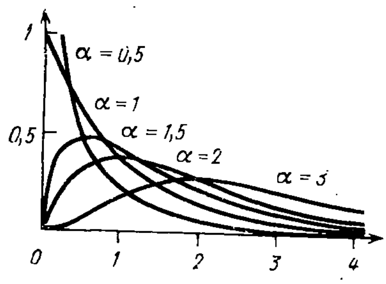

Рис. 17

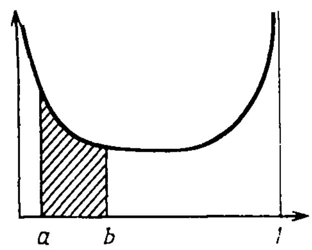

Рис. 18

Допустим, что $\mathbf{P}$ есть произвольное распределение вероятностей на прямой, $F(x)=\mathbf{P}((-\infty, x])$ — его функция распределения, и обозначим $f(x)$ плотность распределения, если она существует. Рассмотрим линейное растяжение (сжатие) $x\to\lambda^{-1}x$ прямой с распределенной на ней по закону $\mathbf{P}$ вероятностной массой. Результатом этого преобразования будет новое распределение вероятностной массы $\mathbf{P}_\lambda$, которое приписывает произвольному промежутку $(a, b]$ значение

$$\mathbf{P}_\lambda((a, b])=\mathbf{P}((\lambda a, \lambda b]).$$

Соответствующие функции и плотности распределений связаны соотношениями

$$F_\lambda(x)=F(\lambda x),\quad f_\lambda(x)=\lambda f(\lambda x).$$

Отметим, что если на прямой взять новый масштаб, приняв точку λ^{-1} за единичную, то распределение вероятностей $\mathbf{P}_\lambda$ в новом масштабе имеет ту же математическую форму, что и распределение $\mathbf{P}$ в старом. С этой точки зрения распределение $\mathbf{P}_\lambda$ есть просто распределение $\mathbf{P}$, изучаемое в другом масштабе.

Масштабные преобразования относятся к числу важных приемов при изучении распределений вероятностей. Так, равномерное распределение на $[0, 1]$ и гамма-распределение в 3.14 были получены предельным переходом от соответствующим образом масштабированных дискретных распределений. На языке случайных величин масштабные преобразования распределений сводятся к умножению случайной величины на постоянную. Если сл. в. X имеет ф. р. $F(x)$, то сл. в. σX при $\sigma>0$ имеет ф. р.

$$\mathbf{P}(\omega:\sigma X(\omega)\leqslant x)=\mathbf{P}(\omega:X(\omega)\leqslant\sigma^{-1}x)=\mathcal{F}(\sigma^{-1}x).$$

Параметр σ наряду с $\lambda=\sigma^{-1}$ также называют масштабным.

3.15. Задача. Пусть $\mathbf{P}^{(h)}((a, b])\to\mathbf{P}((a, b])$, $h\to 0$ при любых $0<a<b$, где распределение вероятностей $\mathbf{P}$ сосредоточено на $[0, \infty)$, непрерывно. Показать, что из этого вытекает, что $\mathbf{P}^{(h)}([0, b])\to\mathbf{P}([0, b])$ при любом $b>0$.

[Для любого $\varepsilon>0$ выберем такое $\delta>0$, что $\mathbf{P}((\delta, \delta^{-1}])>1-\varepsilon$. По условию имеем $\mathbf{P}^{(h)}((\delta, \delta^{-1}])\to\mathbf{P}((\delta, \delta^{-1}])$, $h>0$, и, следовательно, для всех достаточно малых h $\mathbf{P}^{(h)}((\delta, \delta^{-1}])>1-2\varepsilon$, так что $\mathbf{P}^{(h)}((0, \delta])<2\varepsilon$. Отсюда при $\delta<b$

$$\mathbf{P}^{(h)}([0, b])=\mathbf{P}^{(h)}([0, \delta])+\mathbf{P}^{(h)}((\delta, b])\leqslant 3\varepsilon+\mathbf{P}((\delta, b])$$

для всех достаточно малых h. Аналогично

$$\mathbf{P}^{(h)}([0, b])>\mathbf{P}^{(h)}((\delta, b])>\mathbf{P}((\delta, b])-\varepsilon.$$

Объединяя полученные неравенства и пользуясь тем, что

$$\mathbf{P}([0, \delta])<\varepsilon,\quad \mathbf{P}([0, b])-\varepsilon<\mathbf{P}((\delta, b])<\mathbf{P}([0, b]),$$

имеем для всех достаточно малых h

$$\mathbf{P}([0, b])-2\varepsilon<\mathbf{P}^{(h)}([0, b])<3\varepsilon+\mathbf{P}([0, b]).]$$

3.16. Задача. Рассмотрим последовательность дискретных распределений вероятностей, сосредоточенных в точках k/n, $k=0, 1, \ldots, n$:

$$p_n(k/n)=u_{2k}u_{2n-2k},\quad u_{2m}=C_{2m}^m 2^{-2m},$$

$$\mathbf{P}_n((a, b])=\sum_{k=[na]+1}^{[nb]} p_n(k/n).$$

Воспользовавшись *формулой Стирлинга* (см.3.17 и далее), показать, что при $n\to\infty$

$$\mathbf{P}_n((a, b])\to\int_a^b f(x)\,dx,\quad 0\leqslant a<b\leqslant 1,\quad f(x)=\pi^{-1}(x(1-x))^{-1/2}.$$

[По формуле Стирлинга $m!\sim\sqrt{2\pi m}\,e^{-m}m^m$, $m\to\infty$, имеем

$$C_{2m}^m=\frac{2m!}{(m!)^2}\sim\frac{\sqrt{2\pi 2m}\,e^{-2m}(2m)^{2m}}{(\sqrt{2\pi m}\,e^{-m}m^m)^2}=\frac{1}{\sqrt{\pi m}}2^{2m},$$

откуда при $n\to\infty$ равномерно по $k\to\infty$, $n-k\to\infty$ (т. е. равномерно по k, $n-k>\psi(n)$, где $\psi(n)\to\infty$ произвольна)

$$p_n(k/n)\sim\pi^{-1}(k(n-k))^{-1/2}=n^{-1}f(k/n),$$

$$\mathbf{P}_n((a,\ b])\sim\sum_{k=[na]+1}^{[nb]}n^{-1}f(k/n)\to\int_a^b f(x)\,dx,\ 0<a<b<1.$$

Сходимость вероятностей $P_n((a,\ b])$ при $a=0$ или $b=1$ выводится отсюда таким же образом, как это сделано в 3.15.]

Распределение вероятностей с плотностью $f(x)=\pi^{-1}(x(1-x))^{-1/2}$ имеет ф. р. $F(x)=2\pi^{-1}\arcsin\sqrt{x}$, $0\leqslant x\leqslant 1$ (см. 3.8), и называется распределением арксинуса. В схеме симметричного случайного блуждания длительности $2n$ сл. в. σ_{2n} — момент последнего попадания в начальное состояние — имеет распределение $u_{2k}u_{2n-2k}$, $k=0,\ 1,\ \ldots,\ n$ (см. 1.19). Сл. в. $\sigma_{2n}^*=\sigma_{2n}/2n$ имеет распределение вероятностей $\mathbf{P}_n$ из 3.16, так что при $n\to\infty$

$$F_{\sigma_{2n}^*}(x)\to F(x).$$

Сл. в. τ_{2n} — момент первого достижения траекторий длительности $2n$ своего максимума — принимает значения $2k$, $2k+1$ с вероятностями $(1/2)u_{2k}u_{2n-2k}$, $k=1,\ 2,\ \ldots,\ n-1$ (см. 1.18), откуда следует, что при $n\to\infty$

$$F_{\tau_{2n}^*}(x)\to F(x),\ \ \tau_{2n}^*=\tau_{2n}/2n.$$

График арксинус-плотности $f(x)$ представлен на рис. 18. Поскольку вероятностная мера отрезка представляется площадью построенной над ним криволинейной трапеции, то из рисунка заключаем, что наиболее вероятные области отрезка $[0,\ 1]$ примыкают к его концам. К примеру, решая уравнения $F(x)=1/3$, $F(x)=2/3$, находим $x=1/4$, $x=3/4$ соответственно, так что на отрезках $[0,\ 1/4]$, $[1/4,\ 3/4]$, $[3/4,\ 1]$ сосредоточены равные вероятностные массы $1/3$. Таким образом, при больших n около 0,66 всех траекторий достигают в первый раз максимума в первую либо последнюю четверти полного времени блуждания; около 0,33 всех траекторий, начинающихся из 0, не меняют знака на последних трех четвертях времени блуждания. Еще более неожиданно, что на отрезки $[0;\ 0{,}1]$, $[0{,}1;\ 0{,}9]$, $[0{,}9;\ 1]$ приходятся соответственно вероятности $\approx 0{,}2;\ 0{,}6;\ 0{,}2$.

Наконец, отметим, что асимптотическая формула

$$u_{2k}u_{2n-2k}\sim n^{-1}f(k/n),\quad n\to\infty,$$

объясняет, почему огибающие на рис. 8 имеют форму арксинус-плотности.

3.17. Задача. Вывести формулу

$$\ln n!=n\ln n-n+2^{-1}\ln n+\mathrm{const}+o(1),\quad n\to\infty,$$

приближая сумму $\ln 1+\ln 2+\ldots+\ln n$ интегралом $\int\limits_0^n \ln x\,dx$ с помощью соотношения

$$\int\limits_k^{k+1} \ln x\,dx = 2^{-1}(\ln k+\ln(k+1))+O(k^{-2}).$$

$$\Big[\Delta_k = \int\limits_k^{k+1} \ln x\,dx - 2^{-1}(\ln k+\ln(k+1)) =$$

$$= 2^{-1}\int\limits_k^{k+1} (\ln(x/(k+1))+\ln(x/k))\,dx =$$

$$= 2^{-1}\int\limits_k^{k+1} (\ln(1+(x-k-1)/(k+1))+\ln(1+(x-k)/k))\,dx =$$

$$= 2^{-1}\int\limits_0^1 (\ln(1-y/(k+1))+\ln(1+y/k))\,dy =$$

$$= 2^{-1}\int\limits_0^1 (-y/(k+1)+y/k+O(k^{-2}))\,dy = O(k^{-2}).$$

Ввиду выпуклости функции $\ln x$ $\Delta_k>0$, а из предыдущего вытекает, что сходится ряд $\sum\limits_{k=1}^{\infty} \Delta_k = c$. Таким образом,

$$\int\limits_1^n \ln x\,dx = \sum_{k=1}^{n-1}\int\limits_k^{k+1} \ln x\,dx = 2^{-1}\sum_{k=1}^{n-1}(\ln(k+1)+\ln k) +$$

$$+\sum_{k=1}^{n-1}\Delta_k = \sum_{k=1}^{n}\ln k - 2^{-1}\ln n + c + o(1), \quad n\to\infty.$$

Интегрированием по частям находим

$$\int\limits_1^n \ln x\,dx = n\ln n - n + 1,$$

откуда

$$\ln n! = \sum_{k=1}^{n}\ln k = n\ln n - n + 2^{-1}\ln n - c + 1 + o(1), \quad n\to\infty.\Big]$$

Формулу Стирлинга

$$n! \sim \sqrt{2\pi n}\,e^{-n}n^n, \quad n\to\infty,$$

выведем из 3.17, установив, что в формуле

$$n! \sim a\sqrt{n}\, e^{-n} n^n, \quad n \to \infty,$$

постоянная $a = \sqrt{2\pi}$. Используем для этой цели задачу 3.16. Введем обозначение $\alpha_n \asymp \beta_n$, если отношение α_n/β_n заключено между двумя положительными постоянными при всех $n=1, 2, \ldots$. Из 3.17 вытекает, что

$$n! \asymp \sqrt{n}\, e^{-n} n^n.$$

Используя эту формулу вместо формулы Стирлинга в рассуждениях, проведенных в 3.16, получим

$$p_n(k/n) \asymp (k(n-k))^{-1/2}, \quad k=1, \ldots, n-1, \quad n=2, 3, \ldots .$$

Отсюда при любом $0<\delta<1$ и $n \geqslant 1/\delta$ имеем

$$\mathbf{P}_n([0, \delta]) = p_n(0) + \sum_{k=1}^{[n\delta]} p_n(k/n) \asymp \int_0^\delta (x(1-x))^{-1/2}\, dx.$$

В таком случае по заданному $\varepsilon>0$ можно выбрать $\delta>0$ настолько малым, что при $n \geqslant 1/\delta$

$$\mathbf{P}_n([0, \delta] \cup [1-\delta, 1]) = 2\mathbf{P}_n([0, \delta]) < \varepsilon,$$

или

$$\mathbf{P}_n((\delta, 1-\delta]) > 1-\varepsilon.$$

Снова обратимся к решению задачи 3.16, используя вместо формулы Стирлинга формулу $n! \sim a\sqrt{n} e^{-n} n^n$. В результате получим для $0<\alpha<\beta<1$

$$\mathbf{P}_n((\alpha, \beta]) \to \int_\alpha^\beta 2a^{-2} (x(1-x))^{-1/2}\, dx \leqslant 1$$

и, в частности,

$$1-\varepsilon < \mathbf{P}_n((\delta, 1-\delta]) \to \int_\delta^{1-\delta} 2a^{-2} (x(1-x))^{-1/2}\, dx \leqslant 1.$$

Устремляя ε к 0, получаем

$$\int_0^1 2a^{-2} (x(1-x))^{-1/2}\, dx = 1.$$

Но

$$\int_0^1 (x(1-x))^{-1/2}\, dx = 2\arcsin \sqrt{x}\, \Big|_0^1 = \pi,$$

так что $2a^{-2} = \pi^{-1}$, $a = \sqrt{2\pi}$.

3.18. Задача. Показать, что при $n \to \infty$

$$\ln (C_{2n}^{n+k}/C_{2n}^{n}) = -k^2/n + o(1).$$

где остаточный член стремится к нулю равномерно по $k = o(n^{3/4})$ (т. е. равномерно по $k < \psi(n) n^{3/4}$, где $\psi(n) \to 0$ произвольная). [При $0 < k \leqslant n$ имеем

$$C_{2n}^{n+k}/C_{2n}^{n} = \frac{(2n)!\, n!\, n!}{(n+k)!\,(n-k)!\,(2n)!} = \frac{n(n-1)\ldots(n-k+1)}{(n+k)(n+k-1)\ldots(n+1)} =$$

$$= \left(1 - \frac{1}{n}\right) \ldots \left(1 - \frac{k-1}{n}\right) \left(1 + \frac{1}{n}\right)^{-1} \ldots \left(1 + \frac{k}{n}\right)^{-1}.$$

Переходя к логарифмам и используя разложение Тейлора

$$\ln(1+x) = x - x^2/2 + O(1) x^3 \quad x \to 0,$$

получаем при $n \to \infty$ и $k = o(n)$

$$\ln (C_{2n}^{n+k}/C_{2n}^{n}) = \sum_{i=0}^{k-1} \left(\ln\left(1 - \frac{i}{n}\right) - \ln\left(1 + \frac{i+1}{n}\right)\right) =$$

$$= \sum_{i=0}^{k-1} \left(-\frac{i}{n} - \frac{1}{2}\left(\frac{i}{n}\right)^2 - \frac{i+1}{n} + \frac{1}{2}\left(\frac{i+1}{n}\right)^2 + O\left(\frac{i}{n}\right)^3 \right) =$$

$$= -2 \sum_{i=0}^{k-1} \frac{i}{n} - \frac{k}{n} + \frac{1}{2}\left(\frac{k}{n}\right)^2 + O(1) \sum_{i=0}^{k-1} \left(\frac{i}{n}\right)^3 =$$

$$= -\frac{(k-1)k}{n} + o(1) + O(1)\frac{k^4}{n^3} = -\frac{k^2}{n} + o(1) + O(1)\frac{k^4}{n^3}.$$

Выбирая $k = o(n^{3/4})$, получаем требуемое соотношение.]

3.19. Пример (модель броуновского движения). Английский ботаник Р. Броун обнаружил в 1828 г., что частицы пыльцы, взвешенные в воде, совершают непрерывное беспорядочное движение. Молекулярная теория вещества рассматривает это движение как результат «бомбардировки» частицы колеблющимися молекулами воды. Рассмотрим проекцию броуновского движения на ось и опишем ее приближенно моделью симметричного случайного блуждания, полагая, что перемещения блуждающей частицы имеют величину $\pm h$ и происходят в дискретные моменты времени $t = j\delta$, $j = 0, 1, 2, \ldots$. При $h = \delta = 1$ эта модель изучалась в § 1. Сейчас нас интересует предельный вариант этой модели, когда $\delta \to 0$, $h \to 0$. Нетрудно понять, что соотношение между порядками малости δ и h не может быть произвольным, поскольку связь между пройденным путем h и затраченным временем δ — существенная характеристика движения. Примечательно, что порядок малости h относительно δ может быть найден без обращения к физическому эксперименту, а лишь из условия существования предельного варианта дискретной модели. Как будет показано в 3.20, предположения $h/\sqrt{\delta}\to 0$ и $h/\sqrt{\delta}\to\infty$ не приводят к содержательной модели, и единственный разумный вариант — это $h/\sqrt{\delta}\to\sigma$, $0<\sigma<\infty$. Проводимые далее выкладки слегка упрощаются, если положим $h/\sqrt{\delta}=\sigma$. Кроме того, значение σ определяет лишь масштабное преобразование, и мы выберем $\sigma=1$.

Итак, пусть S_n — положение в n-й момент времени блуждающей частицы в модели симметричного случайного блуждания 1.11. Определим случайный процесс с непрерывным временем и кусочно-постоянными траекториями, полагая

$$W_t^{(h)}=hS_{[t/\delta]},\quad h=\sqrt{\delta},$$

где $[a]$ обозначает целую часть числа a. Пусть $n=[t/\delta]$, тогда

$$\mathbf{P}(\omega: W_t^{(h)}(\omega)=kh)=C_n^{(n+k)/2}2^{-n}\equiv p_{n,k},$$

где мы полагаем $C_m^s=0$ при $s<0$, $s>n$, s не целом. Для $n=2m$, используя соотношение (см. 3.16)

$$C_{2m}^m\,2^{-2m}\sim(\pi m)^{-1/2},\quad m\to\infty,$$

и задачу 3.18, имеем при $l=o(m^{3/4})$, $m\to\infty$,

$$p_{2m,\,2l}\sim(\pi m)^{-1/2}\exp(-l^2/m).$$

Для $n=2m+1$, $k=2l+1$, имеем

$$p_{2m+1,\,2l+1}=C_{2m+1}^{m+l+1}2^{-2m-1}=\frac{2m+1}{m+l+1}C_{2m}^{m+l}2^{-2m-1}\leqslant p_{2m,\,2l}.$$

Вероятность

$$\mathbf{P}(\omega: W_t^{(h)}(\omega)\in[a,\ b])=\sum_{\{kh\in[a,b]\}}p_{n,k}$$

при любых фиксированных $a<b$ и $n\to\infty$, очевидно, эквивалентна сумме

$$\frac{1}{2}\sum_{\{kh\in[a,b]\}}(\pi n/2)^{-1/2}\exp(-k^2/(2n)),$$

где множитель 1/2 компенсирует то, что сумма взята по целым k, тогда как $p_{n,k}=0$ при k другой четности, чем n. Подставляя сюда $n=[t/\delta]$, $\delta=h^2$, перепишем сумму в эквивалентном при $h\to 0$ виде:

$$\sum_{\{kh\in[a,b]\}}(2\pi t)^{-1/2}h\exp(-k^2h^2/(2t)).$$

Заметим, что она является интегральной для функции

$$(1/\sqrt{2\pi t})\exp(-x^2/(2t))=(1/\sqrt{t})\,\varphi(x/\sqrt{t}).$$

где функция

$$\varphi(x) = (1/\sqrt{2\pi}) \exp(-x^2/2)$$

— плотность распределения вероятностей (будет показано в 3.21), т. е.

$$\int\limits_{-\infty}^{\infty} \varphi(x)\, dx = 1.$$

Это распределение называют *стандартным нормальным*. Таким образом, при $h \to 0$

$$\mathbf{P}(\omega : W_t^{(h)}(\omega) \in [a, b]) \to \int\limits_a^b (1/\sqrt{t})\, \varphi(x/\sqrt{t})\, dx,$$

т. е. предельное распределение для $W_t^{(h)}$ — масштабное преобразование стандартного нормального.

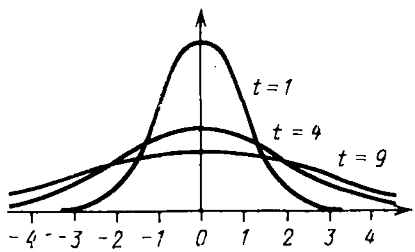

Рис. 19

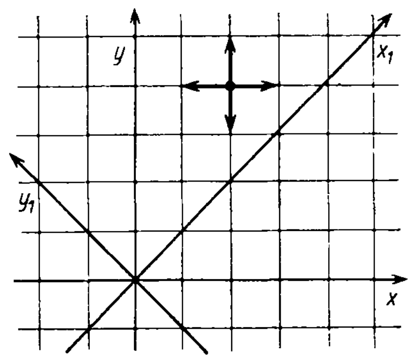

Рис. 20

На рис. 19 представлены графики нормальных плотностей

$$(1/\sqrt{2\pi t}) \exp(-x^2/(2t))$$

для нескольких значений t; с ростом t распределение вероятностей броуновской частицы все больше «расползается» по прямой. Заметим, что вероятность

$$2 \int\limits_{\varepsilon\sqrt{t}}^{\varepsilon^{-1}\sqrt{t}} (1/\sqrt{t})\, \varphi(x/\sqrt{t})\, dx = 2 \int\limits_{\varepsilon}^{\varepsilon^{-1}} \varphi(y)\, dy$$

того, что броуновская частица будет находиться в момент времени t на расстоянии от $\varepsilon\sqrt{t}$ до $\varepsilon^{-1}\sqrt{t}$ от начального положения, может быть сделана сколь угодно близкой к 1 за счет малости ε. Поэтому можно сказать, что при $t\to\infty$ броуновская частица удаляется от начала на величину порядка $\sqrt{t}$.

Отметим, что установленное выше соотношение

$$C_{2n}^{n+k}2^{-2n}\sim C_{2n+1}^{n+k+1}2^{-2n-1}\sim\left(\sqrt{2}/\sqrt{n}\right)\varphi\left(\sqrt{2}\,k/\sqrt{n}\right)$$

объясняет сходство огибающих на рис. 3.

3.20. Задача. В обозначениях 3.19, заметив, что $W_t^{(h)}=hW_t^{(1)}$, показать, что если $\delta\to 0$, $h/\sqrt{\delta}\to 0$, то

$$\mathbf{P}\left(\omega: W_t^{(h)}(\omega)\in[-\alpha,\alpha]\right)\to 1$$

при любом $\alpha>0$, а если $h/\sqrt{\delta}\to\infty$, то

$$\mathbf{P}\left(\omega: W_t^{(h)}(\omega)\in[-\alpha,\alpha]\right)\to 0,$$

т. е. в обоих случаях предельные процессы получаются вырожденные (в одном случае броуновская частица не сдвигается с места, а в другом — мгновенно уходит в бесконечность).

$$[\mathbf{P}\left(\omega: W_t^{(h)}(\omega)\in[-\alpha,\alpha]\right)=\mathbf{P}\left(\omega:(h/\sqrt{\delta})\,W_t^{(\sqrt{\delta})}(\omega)\in[-\alpha,\alpha]\right)=$$
$$=\mathbf{P}\left(\omega: W_t^{(\sqrt{\delta})}(\omega)\in\left[-\alpha\sqrt{\delta}/h,\ \alpha\sqrt{\delta}/h\right]\right).$$

При любом фиксированном $a>0$ полученное выражение оценивается снизу вероятностью

$$\mathbf{P}\left(\omega: W_t^{(\sqrt{\delta})}(\omega)\in[-a,a]\right)$$

при $a<\alpha\sqrt{\delta}/h$, и эта вероятность дает оценку сверху при $a>\alpha\sqrt{\delta}/h$. Учитывая, что предельное значение при $\delta\to 0$ этой вероятности равно

$$\int_{-a}^{a}\left(1/\sqrt{t}\right)\varphi\left(x/\sqrt{t}\right)dx,$$

получим требуемый результат, выбирая a сколь угодно малым в случае $\sqrt{\delta}/h\to\infty$ и сколь угодно большим в случае $\sqrt{\delta}/h\to 0$.]

3.21. Задача. Проверить, что $\varphi(x)=\left(1/\sqrt{2\pi}\right)\exp(-x^2/2)$ является плотностью распределения вероятностей: (I) пользуясь равенством $\Gamma(1/2)=\sqrt{\pi}$; (II) возводя интеграл $\int\varphi(x)\,dx$ в квадрат и переходя в двойном интеграле к полярным координатам.

$$\left[(I)\ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-x^2/2}\,dx=\frac{1}{\sqrt{\pi}}\int_{0}^{\infty}e^{-x^2/2}\left(\frac{x^2}{2}\right)^{-1/2}d\left(\frac{x^2}{2}\right)=\frac{\Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi}},\right.$$

$$(\text{II})\left(\int\limits_{-\infty}^{\infty}\varphi(x)\,dx\right)^2=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}e^{-(x^2+y^2)/2}\,dx\,dy=$$

$$\left.=\frac{1}{2\pi}\int\limits_{0}^{2\pi}d\theta\int\limits_{0}^{\infty}e^{-r^2/2}r\,dr=\int\limits_{0}^{\infty}e^{-r^2/2}r\,dr=1.\right]$$

3.22. Пример (двумерное броуновское движение). Рассмотрим двумерное симметричное случайное блуждание по целочисленной плоской решетке полагая, что перемещения происходят из любой точки (x, y) с целыми координатами в одно из четырех соседних положений:

$$(x+1,\ y),\ (x-1,\ y),\ (x,\ y+1),\ (x,\ y+1).$$

Управляет перемещением частицы повторный случайный выбор из множества $\{(1,\ 0),\ (-1,\ 0),\ (0,\ 1),\ (0,\ -1)\}$. Таким образом, на каждом шаге частица перемещается в одну из четырех соседних точек с равными вероятностями 1/4. Найдем распределение вероятностей положения частицы в момент времени n. Расчет облегчается, если перейти к системе координат x_1, y_1, повернутой относительно x, y на 45° (рис. 20). Узлы целочисленной решетки в системе координат x, y имеют в системе x_1, y_1 координаты, кратные $1/\sqrt{2}$. Подсчитаем число траекторий длительности n, ведущих из начала координат в точку $x_1=k/\sqrt{2}$, $y_1=l/\sqrt{2}$, k, l — целые. Каждый шаг блуждающей частицы означает в системе координат x, y сдвиг на единицу вправо, влево, вверх или вниз, а в системе x_1, y_1 это означает одновременный сдвиг на расстояние $1/\sqrt{2}$ как по горизонтали, так и по вертикали (т. е. параллельно x_1 и y_1). Таким образом, проекции движущейся частицы на x_1, y_1 представляют собой одномерные блуждания с шагом $1/\sqrt{2}$ той же длительности n, как и двумерная траектория. Отсюда число траекторий длительности n, соединяющих точки $(0,\ 0)$ и $(x_1,\ y_1)=\left(k/\sqrt{2},\ l/\sqrt{2}\right)$, равно $N_{n,k}\cdot N_{n,l}$, где $N_{n,i}=C_n^{(n+i)/2}$ — число траекторий одномерного блуждания, оканчивающихся в точке i. Соответствующая вероятность равна $p_{n,k}\cdot p_{n,l}$.

Выберем $h=\sqrt{2\delta}$ и перейдем к блужданию по решетке с шагом h, происходящему в дискретные моменты времени с шагом δ. Как и в 3.19, введем непрерывное время и обозначим $\mathbf{P}_t^{(h)}$ вероятностную меру на плоскости x_1, y_1, полагая

$$\mathbf{P}_t^{(h)}((a_1,\ b_1]\times(a_2,\ b_2])=\sum_{k=[\sqrt{2}a_1/h]+1}^{[\sqrt{2}b_1/h]}\ \sum_{l=[\sqrt{2}a_2/h]+1}^{[\sqrt{2}b_2/h]}p_{n,k}p_{n,l},\quad n=\left[\frac{t}{\delta}\right],$$

равной вероятности попадания частицы в момент t в прямоугольник $(a_1, b_1] \times (a_2, b_2]$ в координатной системе x_1, y_1. При $\delta \to 0$ получаем из 3.19, что

$$\mathbf{P}_t^{(h)}((a_1, b_1] \times (a_2, b_2]) \to \int_{a_1}^{b_1}\int_{a_2}^{b_2} \frac{1}{\sqrt{t}}\, \varphi\left(\frac{x_1}{\sqrt{t}}\right) \frac{1}{\sqrt{t}}\, \varphi\left(\frac{y_1}{\sqrt{t}}\right) dx_1\, dy_1.$$

Предельное распределение вероятностей на плоскости x_1, y_1 равно

$$\mathbf{P}_t((a_1, b_1] \times (a_2, b_2]) = \int_{a_1}^{b_1} \frac{1}{\sqrt{t}}\, \varphi\left(\frac{x_1}{\sqrt{t}}\right) dx_1 \cdot \int_{a_2}^{b_2} \frac{1}{\sqrt{t}}\, \varphi\left(\frac{y_1}{\sqrt{t}}\right) dy_1$$

и, как видно, распадается в произведение одномерных распределений, отвечающих одномерным броуновским движениям. Стягивая прямоугольник $(a_1, b_1] \times (a_2, b_2]$ к точке (x_1, y_1), получаем

$$\mathbf{P}_t((a_1, b_1] \times (a_2, b_2])/((b_1 - a_1)(b_2 - a_2)) \to (2\pi t)^{-1}\, e^{-(x_1^2 + y_1^2)/(2t)},$$

так что вероятностная мера $\mathbf{P}_t$ на плоскости (x_1, y_1) имеет *двумерную плотность распределения* $(2\pi t)^{-1} \exp(-(x_1^2 + y_1^2)/(2t))$. Вероятность любой двумерной области B может быть выражена интегралом от плотности:

$$\mathbf{P}_t(B) = \iint_B (2\pi t)^{-1} \exp(-(x_1^2 + y_1^2)/(2t))\, dx_1\, dy_1,$$

(в предположении, что интеграл по этой области определен).

3.23. Задача. Найти функцию распределения вероятностей сл. в. ρ_t — расстояния броуновской частицы от начала координат в момент t в двумерной модели 3.22.

$$\Big[F_{\rho_t}(u) = \iint_{\{x^2 + y^2 \leqslant u^2\}} (2\pi t)^{-1} \exp(-(x^2 + y^2)/(2t))\, dx\, dy =$$

$$= \int_0^{2\pi} d\theta \int_0^u (2\pi t)^{-1} \exp(-r^2/(2t))\, r\, dr = 1 - \exp(-u^2/(2t)), \quad u \geqslant 0.\Big]$$

Отметим, что вероятность

$$F_{\rho_t}(\varepsilon^{-1}\sqrt{t}) - F_{\rho_t}(\varepsilon\sqrt{t}) = e^{-\varepsilon^2/2} - e^{-\varepsilon^{-2}/2}$$

того, что броуновская частица находится в момент t на расстоянии от $\varepsilon\sqrt{t}$ до $\varepsilon^{-1}\sqrt{t}$ от начала координат, может быть сделана сколь угодно близкой к 1 за счет малости ε. Таким образом, это расстояние имеет с ростом t порядок $\sqrt{t}$.

Абсолютно непрерывное вероятностное распределение в n-мерном пространстве R^n задается с помощью неотрицательной интегрируемой функции $f(x_1, \ldots, x_n)$, такой, что

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = 1,$$

по формуле

$$\mathbf{P}(A)=\int\limits_A\ldots\int f(x_1,\ \ldots,\ x_n)\,dx_1\ldots dx_n,$$

которая определяет меру $\mathbf{P}$ на системе множеств $\mathscr{A}$, для которых интеграл определен и является аддитивной функцией множества. Функция f называется плотностью меры $\mathbf{P}$. В точках непрерывности функции f имеем

$$f(x_1,\ldots,x_n)=(\partial^n/(\partial x_1\ldots\partial x_n))F(x_1,\ldots,x_n),$$

$$F(x_1,\ldots,x_n)=\mathbf{P}((-\infty,\ x_1]\times\ldots\times(-\infty,\ x_n]).$$

На вероятностном пространстве $(R^n,\ \mathscr{A},\ \mathbf{P})$ можно задать непосредственно сл. в. $X_k(\mathbf{x})=x_k$, $\mathbf{x}=(x_1,\ldots,x_n)$, $k=1,\ldots,n$, с совместной ф. р. F и *совместной плотностью распределения* f.

3.24. Задача. Сл. в. $X_1,\ X_2,\ldots,X_n$ — координаты n брошенных наудачу точек на отрезок $[0,\ 1]$ имеют совместную плотность $f(x_1,\ x_2,\ldots,x_n)=1$ при $x_1,\ x_2,\ldots,x_n\in[0,\ 1]$. Найти совместную плотность сл. в. $X_{(1)},\ X_{(2)},\ldots,X_{(n)}$, где $X_{(k)}$ — положение k-й по порядку точки, считая от левого конца отрезка.

[Для любого паралеллепипеда $[a_1,\ b_1]\times[a_2,\ b_2]\times\ldots\times[a_n,\ b_n]$, такого, что $0\leqslant a_1<b_1\leqslant a_2<b_2\leqslant\ldots\leqslant a_n<b_n\leqslant 1$, имеем

$$\mathbf{P}(\omega:X_{(k)}(\omega)\in[a_k,\ b_k],\ k\leqslant n)=\sum\mathbf{P}(\omega:X_{i_k}(\omega)\in[a_k,\ b_k],\ k\leqslant n),$$

где суммирование ведется по всем $n!$ перестановкам $(i_1,\ldots,i_n)$ из чисел $1,\ldots,n$. Продолжая равенство, получаем

$$n!\,(b_1-a_1)\ldots(b_n-a_n)=\int\limits_{a_1}^{b_1}\ldots\int\limits_{a_n}^{b_n}n!\,dx_1\ldots dx_n.$$

Таким образом, для n-мерных параллелепипедов, содержащихся в области $0\leqslant x_1\leqslant x_2\leqslant\ldots\leqslant x_n\leqslant 1$, имеем формулу

$$\mathbf{P}_{X_{(1)},\ldots,X_{(n)}}(A)=\int\limits_A\ldots\int n!\,dx_1\ldots dx_n.$$

Из теории интегрирования по Риману вытекает, что это равенство можно распространить на класс множеств $\mathscr{A}$, n-мерные объемы которых могут быть приближены объемами фигур, разбивающихся на параллелепипеды, со сторонами, параллельными координатным осям. Такая фигура, в частности $\{(x_1,\ldots,x_n):0\leqslant x_1\leqslant x_2\leqslant\ldots\leqslant x_n\leqslant 1\}$, откуда находим, что совместная плотность распределения сл. в. $X_{(1)},\ X_{(2)},\ldots,X_{(n)}$ равна $n!$ в пределах указанного множества и равна нулю вне его.]

3.25. Задача. Сл. в. $X_1,\ldots,X_n$ имеют совместную плотность $f(x_1,\ldots,x_n)$. Найти частные плотности $f_{X_i}(x_i)$, $i=1,\ \ldots,\ n$.

[Полагая для краткости $n=2$, получаем

$$F_{X_1}(u)=\mathbf{P}_{X_1,X_2}((-\infty,\ u]\times(-\infty,\ \infty))=\int\limits_{-\infty}^{u}\Big(\int\limits_{-\infty}^{\infty}f(x_1,\ x_2)\,dx_2\Big)\,dx_1$$

в предположении, что двойной интеграл можно заменить повторным. Следовательно,

$$f_{X_1}(x_1)=\int\limits_{-\infty}^{\infty}f(x_1,\ x_2)\,dx_2.]$$

3.26. Задача. Сл. в. $X_1,\ldots,X_n$ имеют совместную плотность $f(x_1,\ldots,x_n)$. Показать, что совместная плотность сл. в.

$$Y_j=\sum_{i=1}^{n}a_{ij}X_i,\ \ j=1,\ \ldots,\ n,\ \ \det\|a_{ij}\|\neq 0,$$

равна

$$f_{Y_1,\ldots,Y_n}(y_1,\ \ldots,\ y_n)=|\det\|a_{ij}\||^{-1}f(x_1,\ \ldots,\ x_n),\ \ y_j=\sum_{i=1}^{n}a_{ij}x_i,\ \ j\leqslant n,$$

или, полагая $\mathbf{Y}=(Y_1,\ \ldots,\ Y_n)$, $\mathbf{X}=(X_1,\ \ldots,\ X_n)$, $A=\|a_{ij}\|$, $\mathbf{x}=(x_1,\ \ldots,\ x_n)$, $\mathbf{y}=(y_1,\ \ldots,\ y_n)$, формулу преобразования плотности при линейном невырожденном преобразовании случайных величин можно записать в виде

$$f_{\mathbf{Y}}(\mathbf{y})=|\det A|^{-1}f_{\mathbf{X}}(\mathbf{y}A^{-1}).$$

$$\Big[\mathbf{P}(\omega:\mathbf{Y}(\omega)\in B)=\mathbf{P}(\omega:\mathbf{X}(\omega)A\in B)=\mathbf{P}(\omega:\mathbf{X}(\omega)\in BA^{-1})=$$

$$=\int\ldots\int\limits_{\{BA^{-1}\}}f(x_1,\ \ldots,\ x_n)\,dx_1\,\ldots\,dx_n).$$

Производя замену переменных $\mathbf{y}=\mathbf{x}A$, получаем

$$\mathbf{P}(\omega:\mathbf{Y}(\omega)\in B)=\int\ldots\int\limits_{\{B\}}|\det A|^{-1}f(\mathbf{y}A^{-1})\,dy_1\,\ldots\,dy_n.\Big]$$

3.27. Задача. Пусть сл. в. $X_1,\ldots,X_n$ имеют сферически-симметричную совместную плотность: $f_{X_1,\ldots,X_n}(x_1,\ \ldots,\ x_n)=h(x_1^2+\ldots+x_n^2)$. Найти плотность сл. в. $Y_1,\ldots,Y_n$, связанных с $X_1,\ldots,X_n$ ортогональным преобразованием.

[Для ортогональной матрицы A имеем $A^{-1}=A$, $|\det A|=1$, $\|\mathbf{y}A\|=\|\mathbf{y}\|$, где $\|\mathbf{y}\|$ — евклидова норма вектора $\mathbf{y}$, откуда

$$f_{Y_1,\ldots,Y_n}(y_1,\ \ldots,\ y_n)=h(y_1^2+\ldots+y_n^2).]$$

Применяя 3.27 к примеру 3.22, заключаем, что координаты броуновской частицы в любой декартовой системе координат (с общим началом) имеют совместную плотность $(2\pi t)^{-1}\exp(-(x^2+y^2)/2t)$.

§ 4.

НЕЗАВИСИМОСТЬ

Рассматривая два или более вероятностных опыта, можно поставить вопрос о наличии связи между их исходами. Так, если производится повторный случайный выбор без возвращения, то после каждого извлечения состав урны изменяется и результат последующего извлечения находится в зависимости от результатов предшествующих извлечений. Напротив, при выборе с возвращением состав урны и условия проведения опыта остаются неизменными: при любом извлечении имеется одно и то же число равновозможных исходов.

Совместное изучение нескольких опытов осуществляется в теории вероятностей путем построения единой вероятностной модели, включающей все эти опыты. Пусть $\Omega=\{1, 2, \ldots, N\}$ — пространство элементарных событий, соответствующее однократному случайному выбору. Повторному выбору с возвращением объема n отвечает пространство $\Omega^n=\Omega\times\ldots\times\Omega=\{\boldsymbol{\omega}=(\omega_1, \ldots, \omega_n)\}$ всевозможных последовательностей $\boldsymbol{\omega}$ длины n из чисел от 1 до N с равномерным распределением на нем. Отдельные опыты могут быть представлены в этой модели сл. в. $X_k(\boldsymbol{\omega})=\omega_k$, $\boldsymbol{\omega}=(\omega_1, \ldots, \omega_n)$, $k=1, \ldots, n$: элементарному событию ω_k k-го опыта отвечает событие $\{\boldsymbol{\omega} : X_k(\omega)=\omega_k\}$ в модели $(\Omega^n, p(\boldsymbol{\omega})=N^{-n})$. Поставим вопрос о выделении того свойства модели $(\Omega^n, p(\boldsymbol{\omega}))$ или последовательности сл. в. $X_1, \ldots, X_n$, которое соответствует содержательному представлению об отсутствии зависимости между отдельными опытами. С этой целью рассмотрим произвольные функции $Y_k(\boldsymbol{\omega})=g_k(X_k(\boldsymbol{\omega}))=g_k(\omega_k)$ от результатов отдельных опытов. Найдем совместное и частные распределения сл. в. $Y_1, \ldots, Y_n$:

$$p_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)=\sum_{\{\boldsymbol{\omega}: g_k(\omega_k)=y_k, k=1,\ldots,n\}} N^{-n}=$$

$$=\sum_{\{\omega_1: g_1(\omega_1)=y_1\}} \cdots \sum_{\{\omega_n: g_n(\omega_n)=y_n\}} N^{-n},$$

$$p_{Y_k}(y_k)=\sum_{\{\boldsymbol{\omega}: g_k(\omega_k)=y_k\}} N^{-n}=\sum_{\{\omega_k: g_k(w_k)=k_k\}} N^{-}$$

откуда получаем соотношение

$$p_{Y_1,Y_2,\ldots,Y_n}(y_1, y_2, \ldots, y_n)=p_{Y_1}(y_1)\, p_{Y_2}(y_2) \cdots p_{Y_n}(y_n).$$

Поясним, почему это соотношение следует считать формальным выражением отсутствия зависимости между сл. в. $Y_1, \ldots, Y_n$. Во-первых, как мы показали, оно выполняется для случайных величин $g_k(\omega_k)$, зависящих от результатов различных испытаний в схеме повторного выбора с возвращением, т. е. таких случайных величин, которые с содержательной точки зрения надо полагать независимыми. Во-вторых, рассмотрим некоторые сл. в. $\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n$, для которых

$$p_{\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n}(y_1, y_2, \ldots, y_n) = p_{\widehat{Y}_1}(y_1)\, p_{\widehat{Y}_2}(y_2) \ldots p_{\widehat{Y}_n}(y_n).$$

Предположим дополнительно, что сл. в. $\widehat{Y}_i$, $i=1, 2, \ldots, n$, имеют конечное число значений, которые они принимают с рациональными вероятностями. В таком случае легко определить сл. в. Y_i, $i=1, 2, \ldots, n$, в схеме повторного выбора с возвращением, как это было сделано выше, так что

$$p_{Y_1, \ldots, Y_n}(y_1, y_2, \ldots, y_n) = p_{\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n}(y_1, y_2, \ldots, y_n).$$

Это означает, что наборы сл. в. $(Y_1, Y_2, \ldots, Y_n)$ и $(\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n)$ представляют собой эквивалентные объекты теории. Следовательно, сл. в. $\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n$ «устроены» так же, как если бы они зависели от различных испытаний в схеме повторного случайного выбора с возвращением.

Переходя к произвольным дискретным сл. в. $X_i(\omega)$, $i=1, 2, \ldots, n$, назовем их *независимыми,* если их совместное распределение вероятностей распадается в произведение частных распределений:

$$p_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = p_{X_1}(x_1)\, p_{X_2}(x_2) \ldots p_{X_n}(x_n).$$

Независимые сл. в. $X_1, X_2, \ldots, X_n$ порождают вероятностное пространство $(\mathscr{X}, p(\mathbf{x}))$ специального вида: $\mathscr{X} = \mathscr{X}_1 \times \ldots \times \mathscr{X}_n$ $p(\mathbf{x}) = p_{X_1}(x_1) \ldots p_{X_n}(x_n)$, $\mathbf{x} = (x_1, \ldots, x_n)$, где $p_{X_i}(x_i)$ представляет собой распределение вероятностей на $\mathscr{X}_i$, $i=1, \ldots, n$. Пространство $(\mathscr{X}, p(\mathbf{x}))$ называют *прямым произведением* пространств $(\mathscr{X}_i, p_{X_i}(x_i))$, $i=1, 2, \ldots, n$.

Имея произвольные вероятностные пространства $(\mathscr{X}_i, p(x_i))$, $i=1, 2, \ldots, n$, $(\mathscr{X}_i \subseteq R)$, можно образовать их прямое произведение $(\mathscr{X}, p(\mathbf{x}))$ и задать на нем непосредственно сл. в. $X_i(\mathbf{x}) = x_i$, $\mathbf{x} = (x_1, \ldots, x_n)$, $i=1, \ldots, n$, которые оказываются независимыми и имеют заданные частные распределения $p_i(x_i)$, $i=1, \ldots, n$. Прямое произведение точно так же образуется из дискретных вероятностных пространств $(\Omega_i, p_i(\omega_i))$ с множествами Ω_i произвольной природы. Легко убедиться, как и выше, что любые функции $X_i(\boldsymbol{\omega}) = g_i(\omega_i)$, $i=1, \ldots, n$, $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)$ будут при этом независимыми случайными величинами. Вероятностное пространство $(\Omega_1 \times \ldots \times \Omega_n, p_1(\omega_1) \ldots p_n(\omega_n))$ назовем пространством независимых испытаний. В случае, когда все $(\Omega_i, p_i(\omega_i))$ совпадают между собой, будем говорить о *повторных независимых испытаниях.*

Примером повторных независимых испытаний служат испытания Бернулли, где $\Omega_i = \{0, 1\}$. В 2.17 было показано, что независимыми являются сл. в. T_i, $i=1, 2, \ldots, r$, равные количеству неудач между последовательными успехами в схеме бесконечной последовательности испытаний Бернулли. Напротив, сл. в. V_i, $i=1, 2, \ldots, r$, в 2.18, равные длинам последовательных серий из успехов или неудач в бесконечной последовательности испытаний Бернулли с вероятностью успеха $p \neq 1/2$, не являются независимыми. В самом деле, расписав равенство $p_{V_1, V_2}(1, 1) = p_{V_1}(1) \cdot p_{V_2}(1)$

$$(1-p)p(1-p)+p(1-p)p=((1-p)p+p(1-p))(p^2+(1-p)^2),$$

находим из него $p=1/2$ (при $0<p<1$). С другой стороны, подставляя в формулу совместного распределения вероятностей сл. в. $V_1, V_2, \ldots, V_r$ значение $p=1/2$, легко убеждаемся, что оно распадается в произведение частных распределений. Наличие зависимости между сл. в. $V_1, V_2, \ldots$ при $p \neq 1/2$ наиболее ясно выражено при p, очень близких к 0 либо к 1. Скажем, при p, близких к 0, вероятность того, что первая серия будет состоять из неудач и иметь длину k или больше, равна

$$(1-p)^k p + (1-p)^{k+1} p + \ldots = (1-p)^k$$

и остается близкой к 1 при k больших, но таких, что kp мало. Таким образом, если наблюдается длинная первая серия, то скорее всего она состоит из неудач, так что вторая серия будет состоять из успехов и будет короткой. Если первая серия короткая, то вторая серия скорее всего будет длинная.

4.1. Задача. Сл. в. $X_1, X_2, \ldots, X_n$ имеют совместное распределение вида $p_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = h_1(x_1) h_2(x_2) \ldots h_n(x_n)$, где h_i, $i=1, 2, \ldots, n$, — некоторые неотрицательные функции. Показать, что $X_1, X_2, \ldots, X_n$ независимы.

[Рассмотрим для краткости $n=2$. Имеем

$$p_{X_1}(x_1) = \sum_{\{x_2\}} h_1(x_1) h_2(x_2) = h_1(x_1) c_2, \quad p_{X_2}(x_2) = c_1 h_2(x_2),$$

$$c_2 = \sum_{\{x_2\}} h_2(x_2), \quad c_1 = \sum_{\{x_1\}} h_1(x_1), \quad c_1 c_2 = 1,$$

откуда

$$p_{X_1, X_2}(x_1, x_2) = h_1(x_1) h_2(x_2) = (h_1(x_1) c_2)(h_2(x_2) c_1) = p_{X_1}(x_1) p_{X_2}(x_2).]$$

4.2. Задача. Даны независимые сл. в. $X_1, X_2, \ldots, X_n$. Показать, что любое подмножество $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ также состоит из независимых величин.

[Так как перестановка случайных величин, очевидно, не влияет на свойство независимости, то, перенумеровав величины, сведем вопрос к независимости подмножества $X_1, X_2, \ldots, X_k$. По индукции, переходя от n к $n-1$, получаем с использованием 2.8

$$p_{X_1,\ldots,X_{n-1}}(x_1, \ldots, x_{n-1}) = \sum_{\{x_n\}} p_{X_1}(x_1) \ldots p_{X_n}(x_n) =$$

$$= p_{X_1}(x_1) \ldots p_{X_{n-1}}(x_{n-1}) \sum_{\{x_n\}} p_{X_n}(x_n) = p_{X_1}(x_1) \ldots p_{X_{n-1}}(x_{n-1}).]$$

4.3. З а д а ч а. Даны независимые сл. в. $X_1, \ldots, X_n, X_{n+1}, \ldots$ $\ldots, X_{n+m}$. Показать, что сл. в. $U = g(X_1, \ldots, X_n)$ и $V = h(X_{n+1}, \ldots$ $\ldots, X_{n+m})$ также независимы.

[Для краткости записи положим $n = m = 1$. Имеем

$$p_{U,V}(u, v) = \mathbf{P}(\omega : g(X_1(\omega)) = u,\ h(X_2(\omega)) = v) =$$

$$= \sum_{\{x_1, x_2 : g(x_1) = u, h(x_2) = v\}} \mathbf{P}(\omega : X_1(\omega) = x_1,\ X_2(\omega) = x_2) =$$

$$= \sum_{\{x_1 : g(x_1) = u\}} \mathbf{P}(\omega : X_1(\omega) = x_1) \sum_{\{x_2 : h(x_2) = v\}} \mathbf{P}(\omega : X_2(\omega) = x_2) = p_U(u)\, p_V(v).]$$

4.4. З а д а ч а. Пусть сл. в. X, Y независимы. Вывести формулу распределения вероятностей сл. в. $U = X + Y$:

$$p_U(u) = \sum_{\{x, y : x + y = u\}} p_X(x)\, p_Y(y) = \sum_{\{x\}} p_X(x)\, p_Y(u - x).$$

$$\left[p_U(u) = \mathbf{P}(\omega : X(\omega) + Y(\omega) = u) = \sum_{\{x, y : x + y = u\}} \mathbf{P}(\omega : X(\omega) = x,\ Y(\omega) = y.\right]$$

4.5. З а д а ч а. Сл. в. U, V независимы и имеют биномиальные распределения:

$$p_U(k) = C_n^k p^k (1 - p)^{n-k},\quad p_V(l) = C_m^l p^l (1 - p)^{n-l}.$$

Найти распределение вероятностей сл. в. $U + V$: (I) пользуясь 4.4; (II) рассматривая последовательность $n + m$ испытаний Бернулли с вероятностью успеха p.

$$\left[(\mathrm{I})\ p_{U+V}(k) = \sum_{\{i\}} C_n^i p^i (1 - p)^{n-i} C_m^{k-i} p^{k-i} (1 - p)^{m-k+i} =\right.$$

$$= p^k (1 - p)^{n+m-k} \sum_{\{i\}} C_n^i C_m^{k-i} = C_{n+m}^k p^k (1 - p)^{n+m-k},$$

где на последнем шаге учтено тождество (см. 1.28)

$$\sum_{\{i\}} C_n^i C_m^{k-i} / C_{n+m}^k = 1.$$

(II) Обозначив $\widehat{U}$, $\widehat{V}$ число успехов в первых n и последующих m испытаниях в последовательности $n + m$ испытаний Бернулли, получаем, что $\widehat{U}$ и $\widehat{V}$ независимы и имеют биномиальные распределения с параметрами n, p и m, p, а $\widehat{U} + \widehat{V}$ имеет биномиальное распределение с параметрами $n+m$, p, так как представляет собой число успехов в последовательности $n+m$ испытаний Бернулли.]

4.6. Задача. Сл. в U и V независимы и имеют отрицательные биномиальные распределения с одинаковым значением параметра p. Показать, что $U+V$ имеет также отрицательное биномиальное распределение, рассмотрев пару сл. в. $\widehat{U}$, $\widehat{V}$, определенных в схеме бесконечной последовательности испытаний Бернулли и имеющих то же совместное распределение вероятностей, что и пара U, V (см. 2.17). Вывести отсюда, что сумма n независимых геометрически распределенных с параметром p случайных величин имеет отрицательное биномиальное распределение с параметрами n, p.

[В схеме бесконечной последовательности испытаний Бернулли определим сл. в. $\widehat{U}$, равную числу неудач до r-го успеха, и сл. в. $\widehat{V}$, равную числу неудач между r-м и $(r+s)$-м успехами. Тогда $\widehat{U}+\widehat{V}$ представляет собой число неудач до $(r+s)$-го успеха и потому (см. 2.17)

$$p_{\widehat{U}+\widehat{V}}(k)=C^k_{r+s+k-1}p^{r+s}(1-p)^k.$$

Совместное распределение вероятностей $\widehat{U}$ и $\widehat{V}$ равно

$$p_{\widehat{U},\widehat{V}}(i,\ j)=C^i_{r+i-1}p^i(1-p)^r C^j_{s+j-1}p^j(1-p)^s=p_{\widehat{U}}(i)\,p_{\widehat{V}}(j),$$

т. е. сл. в. $\widehat{U}$, $\widehat{V}$ независимы и имеют отрицательные биномиальные распределения, а их сумма также отрицательная биномиальная.]

4.7. Задача. Сл. в. X, Y независимы и имеют пуассоновские распределения вероятностей с параметрами λ, μ:

$$p_X(k)=(\lambda^k/k!)e^{-\lambda},\quad p_Y(l)=(\mu^l/l!)e^{-\mu},\quad k,\ l=0,\ 1,\ 2,\ \dots .$$

Показать, что их сумма $X+Y$ имеет пуассоновское распределение с параметром $\lambda+\mu$.

$$\Big[p_{X+Y}(k)=\sum_{i=0}^{k}(\lambda^i/i!)\,e^{-\lambda}\,(\mu^{k-i}/(k-i)!)\,e^{-\mu}=$$

$$=\frac{1}{k!}e^{-\lambda-\mu}\sum_{i=0}^{k}\frac{k!}{i!\,(k-i)!}\lambda^i\mu^{k-i}=((\lambda+\mu)^k/k!)\,e^{-\lambda-\mu},$$

где на последнем шаге использована формула бинома Ньютона (см. ниже).]

Докажем формулу *бинома Ньютона*

$$(a+b)^n=\sum_{k=0}^{n}C_n^k a^k b^{n-k},\quad n=1,\ 2,\ \dots,$$

или

$$(1+t)^n=\sum_{k=0}^{n} C_n^k t^k.$$

Раскроем скобки в произведении

$$(1+t_1)(1+t_2)\ldots(1+t_n)=\sum_{\{\delta_1,\delta_2,\ldots,\delta_n\}} t_1^{\delta_1} t_2^{\delta_2}\ldots t_n^{\delta_n},$$

где суммирование ведется по всем наборам $\delta_1, \delta_2, \ldots, \delta_n$ из 0 и 1. Полагая $t_1=t_2=\ldots=t_n=t$, получаем

$$(1+t)^n=\sum_{\{\delta_1,\delta_2,\ldots,\delta_n\}} t^{\delta_1+\delta_2+\ldots+\delta_n}=$$

$$=\sum_{k=0}^{n}\sum_{\{\delta_1,\delta_2,\ldots,\delta_n:\delta_1+\ldots+\delta_n=k\}} t^{\delta_1+\ldots+\delta_n}=\sum_{k=0}^{n} C_n^k t^k.$$

В комбинаторике и теории вероятностей существенную роль играет аппарат производящих функций: степенной ряд

$$\mathcal{A}(t)=a_0+a_1t+a_2t^2+\ldots$$

называют *производящей функцией* (п. ф.) *последовательности* a_0, $a_1, a_2, \ldots$; п. ф. конечной последовательности $a_0, a_1, \ldots, a_n$ можно определить, положив $a_{n+1}=a_{n+2}=\ldots=0$. Предполагая, что ряд для $\mathcal{A}(t)$ имеет ненулевой радиус сходимости, запишем выражения для членов последовательности через производные от $\mathcal{A}(t)$ в нуле:

$$a_n=\mathcal{A}^{(n)}(0)/n!,\quad n=0,\ 1,\ 2,\ldots,\ \mathcal{A}^{(0)}(t)\equiv\mathcal{A}(t),$$

так что последовательность восстанавливается по своей производящей функции. Для последовательности $p_0, p_1, p_2, \ldots$, задающей распределение вероятностей ($0\leqslant p_k\leqslant 1$, $p_0+p_1+p_2+\ldots=1$), п. ф. $\mathcal{P}(t)=p_0+p_1t+\ldots$ называется *вероятностной*. Ряд для $\mathcal{P}(t)$ сходится при $|t|\leqslant 1$. *Производящей функцией сл. в.* X с целыми неотрицательными значениями называют п. ф. $\mathcal{P}_X(t)$, отвечающую распределению вероятностей $p_X(k)$, $k=0, 1, 2, \ldots$ сл. в. X.

Наиболее важные применения производящих функций связаны со следующим их свойством: произведение $\mathcal{A}(t)\mathcal{B}(t)$ является п. ф. последовательности $c_n=a_0b_n+a_1b_{n-1}+\ldots+a_nb_0$, $n=0, 1, 2, \ldots$, где $\mathcal{A}(t)$, $\mathcal{B}(t)$ — п. ф. последовательностей $a_0, a_1, a_2, \ldots$ и $b_0, b_1, b_2, \ldots$, соответственно. В самом деле,

$$\mathcal{A}(t)\mathcal{B}(t)=\sum_{n=0}^{\infty} a_n t^n \sum_{m=0}^{\infty} b_m t^m=\sum_{k=0}^{\infty} t^k \sum_{k=0}^{n} a_k b_{n-k}.$$

Для независимых сл. в. X и Y имеем отсюда и из 4.4, что

$$\mathcal{P}_X(t)\mathcal{P}_Y(t)=\mathcal{P}_{X+Y}(t).$$

4.8. З а д а ч а. Доказать тождество $\sum\limits_{\{i\}} C_n^i C_m^{k-i} = C_{n+m}^k$, используя п. ф.

$$\Bigl[\sum_{\{k\}} C_{n+m}^k t^k = (1+t)^{n+m} = (1+t)^n (1+t)^m = \sum_{\{i\}} C_n^i t^i \times$$

$$\times \sum_{\{j\}} C_m^j t^j = \sum_{\{k\}} t^k \sum_{\{i\}} C_n^i C_m^{k-i}.\Bigr]$$

4.9. З а д а ч а. Дифференцируя тождество $1+t+t^2+\ldots=(1-t)^{-1}$, $|t|<1$, вывести формулу

$$(1-t)^{-r} = \sum_{k=0}^{\infty} C_{k+r-1}^k t^k = \sum_{k=0}^{\infty} C_{k+r-1}^{r-1} t^k.$$

Доказать тождество

$$\sum_{\{k\}} C_{k+r-1}^k C_{l-k+s-1}^{l-k} = C_{l+r+s-1}^l .$$

$$\Bigl[(r-1)!\,(1-t)^{-r} = \sum_{n=r-1}^{\infty} n(n-1)\ldots(n-r+2)\, t^{n-r+1} =$$

$$= \sum_{k=0}^{\infty} (k+r-1)\ldots(k+1)\, t^k = \sum_{k=0}^{\infty} ((k+r-1)!/k!)\, t^k.$$

$$\sum_{l=0}^{\infty} C_{l+r+s-1}^l t^l = (1-t)^{-r} (1-t)^{-s} = \sum_{k=0}^{\infty} C_{k+r-1}^k t^k \times$$

$$\times \sum_{j=0}^{\infty} C_{j+s-1}^j t^j = \sum_{l=0}^{\infty} t^l \sum_{\{k\}} C_{k+r-1}^k C_{l-k+s-1}^{l-k}.\Bigr]$$

4.10. З а д а ч а. Вычислить п. ф. биномиального, отрицательного биномиального и пуассоновского распределений.

$$\Bigl[\sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} t^k = \sum_{k=0}^{n} C_n^k (pt)^k (1-p)^{n-k} = (pt+1-p)^n,$$

$$\sum_{k=0}^{\infty} C_{k+r-1}^k (1-p)^k p^r t^k = p^r (1-(1-p)\,t)^{-r}.$$

(Отсюда понятно, как возникли наименования биномиального и отрицательного биномиального распределений.) Наконец

$$\sum_{k=0}^{\infty} (\lambda^k/k!)\, e^{-\lambda} t^k = e^{-\lambda} \sum_{k=0}^{\infty} (\lambda t)^k/k! = e^{-\lambda+\lambda t}.\Bigr]$$

4.11. Задача. Решить 4.5—4.7 с помощью производящих функций.

$$[(\mathrm{I})\quad \mathcal{P}_{U+V}(t)=\mathcal{P}_U(t)\,\mathcal{P}_V(t)=(pt+1-p)^{k+l},$$

что ввиду 4.10 является п. ф. биномиального распределения, и т. д.]

От понятия независимости случайных величин перейдем к независимости событий. События A и B следует считать независимыми, если то, что произошло или нет одно из них, не влияет на то, произошло или нет другое, т. е. понятие независимости событий сводится к независимости сл. в. I_A, I_B — индикаторов событий A, B ($I_A(\omega)=1$ при $\omega\in A$, $I_A(\omega)=0$ при $\omega\overline{\in}A$). События $A_1, A_2, \ldots, A_n$ называются *независимыми* (в совокупности), если независимы их индикаторы $I_{A_k}(\omega)$, $k=1, \ldots, n$:

$$\mathbf{P}(\omega: I_{A_k}(\omega)=\delta_k,\ k=1,\ \ldots,\ n)=\prod_{k=1}^{n}\mathbf{P}(\omega: I_{A_k}(\omega)=\delta_k)$$

для любого набора $\delta_1, \ldots, \delta_n$ из 0 и 1. Вводя обозначение $A^1=\bar{A}$, $A^0=A$, перепишем условие независимости событий в виде

$$\mathbf{P}(A_1^{\delta_1}A_2^{\delta_2}\ldots A_n^{\delta_n})=\mathbf{P}(A_1^{\delta_1})\,\mathbf{P}(A_2^{\delta_2})\ldots\mathbf{P}(A_n^{\delta_n})$$

для всех наборов $\delta_1, \delta_2, \ldots, \delta_n$.

4.12. Задача. Показать, что для независимости событий A, B (индикаторов I_A, I_B) достаточно выполнения любого из равенств $\mathbf{P}(AB)=\mathbf{P}(A)\mathbf{P}(B)$, $\mathbf{P}(\bar{A}B)=\mathbf{P}(\bar{A})\mathbf{P}(B)$, $\mathbf{P}(A\bar{B})=\mathbf{P}(A)\mathbf{P}(\bar{B})$, $\mathbf{P}(\bar{A}\bar{B})=\mathbf{P}(\bar{A})\mathbf{P}(\bar{B})$.

[Необходимо показать, что любое из выписанных равенств влечет все остальные. Используя разбиение A, $\bar{A}$, запишем: $\mathbf{P}(B)=\mathbf{P}(AB)+\mathbf{P}(\bar{A}B)$ или $\mathbf{P}(\bar{A}B)=\mathbf{P}(B)-\mathbf{P}(AB)$. Предположив, что $\mathbf{P}(AB)=\mathbf{P}(A)\mathbf{P}(B)$, получаем отсюда $\mathbf{P}(\bar{A}B)=\mathbf{P}(B)-\mathbf{P}(A)\mathbf{P}(B)=(1-\mathbf{P}(A))\mathbf{P}(B)=\mathbf{P}(\bar{A})\mathbf{P}(B)$. Точно так же выводятся и остальные равенства.]

4.13. Задача. Индикаторные сл. в. X_1, X_2 независимы и $\mathbf{P}(\omega: X_2(\omega)=1)=1/2$. Образуем сл. в. $X_1\oplus X_2$ — сумму по модулю 2 сл. в. X_1 и X_2 ($1\oplus1=0\oplus0=0$, $0\oplus1=1\oplus0=1$). Показать, что любые две из сл. в. X_1, X_2, $X_1\oplus X_2$ независимы (в то время как все три связаны функциональной зависимостью $X_1\oplus X_2\oplus(X_1\oplus X_2)=0$).

$$[\mathbf{P}(\omega: X_1(\omega)=1,\ X_1(\omega)\oplus X_2(\omega)=1)=\mathbf{P}(\omega: X_1(\omega)=1,\ X_2(\omega)=0)=$$
$$=\mathbf{P}(\omega: X_1(\omega)=1)\,\mathbf{P}(\omega: X_2(\omega)=0)=\mathbf{P}(\omega: X_1(\omega)=1)\cdot 1/2,$$
$$\mathbf{P}(\omega: X_1(\omega)\oplus X_2(\omega)=1=\mathbf{P}(\omega: X_1(\omega)=0)\cdot1/2+$$
$$+\mathbf{P}(\omega: X_1(\omega)=1)\cdot1/2=1/2.$$

Следовательно,

$$\mathbf{P}(\omega: X_1(\omega)=1,\ X_1(\omega)\oplus X_2(\omega)=1)=\mathbf{P}(\omega: X_1(\omega)=1)\,\mathbf{P}(\omega: X_1(\omega)\oplus$$
$$\oplus X_2(\omega)=1).$$

Поскольку сл. в. X_1, $X_1 \oplus X_2$ индикаторные, заключаем (см. 4.12), что они независимы.]

4.14. Задача. События $A_1, A_2, \ldots, A_n$ независимы, $\mathbf{P}(A_i) = p_i$, $i=1, \ldots, n$. Найти вероятность того, что: (I) ни одно из этих событий не произошло; (II) произошло ровно одно событие; (III) произошло ровно k событий при дополнительном предположении $p_1 = \ldots = p_n = p$; (IV) произошло хотя бы одно событие.

[(I) $\mathbf{P}(\bar{A}_1 \ldots \bar{A}_n) = (1-p_1) \ldots (1-p_n)$. Используя представление события (II) в виде

$$\bigcup_{k=1}^{n} \bar{A}_1 \ldots \bar{A}_{k-1} A_k \bar{A}_{k+1} \ldots \bar{A}_n,$$

находим его вероятность

$$\sum_{k=1}^{n} (1-p_1) \ldots (1-p_{k-1}) p_k (1-p_{k+1}) \ldots (1-p_n).$$

Событие (III) имеет вероятность $C_n^k p^k (1-p)^{n-k}$. Наконец,

$$\text{(IV)}\ \mathbf{P}\Big(\bigcup_{k=1}^{n} A_k\Big) = 1 - \mathbf{P}\Big(\bigcap_{k=1}^{n} \bar{A}_k\Big) = 1 - (1-p_1) \ldots (1-p_n).]$$

Модель повторных независимых испытаний $(\mathscr{X}^n, p(x_1) \ldots p(x_n))$ допускает обобщение на случай бесконечной последовательности испытаний, как это было проделано в 2.15 для испытаний Бернулли: $\mathscr{X} = \{0, 1\}$, $p(1) = p$, $p(0) = 1-p$. Пространство элементарных событий $\mathscr{X}^\infty$ состоит из бесконечных последовательностей $\mathbf{x} = (x_1, x_2, \ldots)$, $x_k \in \mathscr{X}$, $k=1, 2, \ldots$, вероятностная мера $\mathbf{P}$ определена на алгебре множеств $\mathscr{A}$, состоящей из подмножеств $A \subseteq \mathscr{X}$ вида

$$A = \{\mathbf{x} : (x_1, \ldots, x_n) \in A^{(n)}\}$$

при всевозможных n и $A^{(n)} \subseteq \mathscr{X}^n$. Формула

$$\mathbf{P}(A) = \sum_{(x_1, \ldots, x_n) \in A^{(n)}} p(x_1) \ldots p(x_n)$$

задает корректно вероятностную функцию $\mathbf{P}(A)$ на $\mathscr{A}$, $\mathbf{P}(\varnothing) = 0$, $\mathbf{P}\mathscr{X}^\infty) = 1$, $\mathbf{P}(A)$ аддитивна на $\mathscr{A}$. Доказательство этих утверждений не отличается от проведенного в 2.15, 2.16 для частного случая испытаний Бернулли. Как и в 2.15, остается в силе замечание, что алгебра множеств $\mathscr{A}$ недостаточна для описания многих интересных свойств бесконечной последовательности испытаний.

Отметим, что независимость событий и дискретных случайных величин для произвольного вероятностного пространства определяется такими же соотношениями, как и в случае дискретного пространства.

По аналогии с дискретным случаем определяется непрерывная модель независимых испытаний. Пусть $f_k(x)$, $k=1, \ldots, n$, — произвольные одномерные плотности распределений,

$$\mathbf{P}(A)=\int \ldots \int_A f_1(x_1) \ldots f_n(x_n)\, dx_1 \ldots dx_n$$

— вероятностная мера с плотностью $f_1(x_1) \ldots f_n(x_n)$, определенная на системе подмножеств $\mathscr{A}$, для которой интеграл определен и является аддитивной функцией множества. Вероятностное пространство $(R^n, \mathscr{A}, \mathbf{P})$ называется моделью независимых испытаний; испытание с номером k в этой модели имеет плотность распределения вероятностей $f_k(x)$, $k=1, \ldots, n$. При $f_k(x)=f(x)$, $k=1, \ldots, n$, $(R^n, \mathscr{A}, \mathbf{P})$ назовем повторными независимыми испытаниями с плотностью распределения отдельного испытания $f(x)$. Пространство $(R^\infty, \mathscr{A}, \mathbf{P})$ для описания бесконечной последовательности повторных независимых испытаний с плотностью $f(x)$ вводится по той же схеме, что и в дискретном случае.

Сл. в. $X_1, X_2, \ldots, X_n$, заданные непосредственно в модели независимых испытаний $(R^n, \mathscr{A}, \mathbf{P}): X_i(\mathbf{x})=x_i$, $\mathbf{x}=(x_1, \ldots, x_n)$, $i=1, \ldots$ $\ldots, n$, определяются результатами независимых испытаний, и их следует назвать независимыми. Однако, как и в дискретной модели, свойство независимости случайных величин целесообразно выразить через их распределение вероятностей. Непрерывные сл. в. $X_i(\omega)$, $i=1, \ldots, n$, заданные на некотором вероятностном пространстве $(\Omega, \mathscr{B}, \mathbf{P})$, назовем *независимыми,* если их совместная плотность распадается в произведение частных:

$$f_{X_1, \ldots, X_n}(x_1, \ldots, x_n)=f_{X_1}(x_1) \ldots f_{X_n}(x_n).$$

Бесконечная последовательность сл. в. $X_1, X_2, \ldots$ называется *независимой,* если независима любая конечная подпоследовательность $X_1, \ldots, X_n$.

4.15. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и имеют стандартные нормальные плотности $\varphi(x_i)=(1/\sqrt{2\pi})\exp(-x_i^2/2)$. Показать, что сл. в.

$$Y_j=\sum_{i=1}^{n} a_{ij} X_i, \quad j=1, \ldots, n,$$

где матрица $\|a_{ij}\|$ ортогональна, также независимы и распределены по стандартному нормальному закону.

[Применить 3.26 и 3.27.]

4.16. Задача. Сл. в. X_1, X_2 независимы и имеют стандартные нормальные плотности, сл. в. R, Φ — полярные координаты точки с декартовыми координатами $X_1, X_2: X_1=R\cos\Phi$, $X_2=R\sin\Phi$. Показать, что R и Φ независимы.

[Для подмножества $B \subseteq (0, \infty) \times [0, 2\pi)$ запишем

$$\mathbf{P}_{R,\Phi}(B) = \iint_{\widetilde{B}} (1/(2\pi)) \exp(-(x_1^2 + x_2^2)/2)\, dx_1\, dx_2,$$

$$\widetilde{B} = \{(x_1, x_2) : x_1 = r\cos\varphi,\ y_1 = r\sin\varphi,\ (r, \varphi) \in B\}.$$

Переходя в интеграле к полярным координатам и вычисляя якобиан $|\partial(x_1, x_2)/\partial(r, \varphi)| = r$, получаем

$$\mathbf{P}_{R,\Phi}(B) = \iint_{B} (1/(2\pi)) \exp(-r^2/2)\, r\, dr\, d\varphi,$$

откуда находим совместную плотность сл. в. R, Φ

$$f_{R,\Phi}(r, \varphi) = (1/(2\pi)) \exp(-r^2/2) r,\ r > 0,\ 0 \leqslant \varphi < 2\pi.$$

Вычисляя частные плотности (см. 3.25)

$$f_R(r) = \int_0^{2\pi} f_{R,\Phi}(r, \varphi)\, d\varphi = re^{-r^2/2},\ f_\Phi(\varphi) = 1/(2\pi),\ r > 0,\ 0 \leqslant \varphi < 2\pi,$$

получаем требуемый результат.]

4.17. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и одинаково распределены с экспоненциальной плотностью $f(x) = e^{-x}$, $x > 0$. Введем сл. в. $X_{(k)}$, $k = 1, \ldots, n$, называемые *порядковыми статистиками:* $X_{(1)} = \min\{X_i,\ i = 1, \ldots, n\}$, $X_{(n)} = \max\{X_i,\ i = 1, \ldots, n\}$, $X_{(k)}$ равна k-му по величине в порядке возрастания среди значений $X_1, \ldots, X_n$. Показать, что сл. в. $Y_i = (n - i + 1)\ (X_{(i)} - X_{(i-1)})$, $i = 1, \ldots, n$, $X_{(0)} = 0$, независимы и имеют одинаковые частные плотности $f(y) = e^{-y}$, $y > 0$.

[Для любой перестановки $i_1, i_2, \ldots, i_n$ из чисел $1, 2, \ldots, n$ сл. в. $X_{i_1}, X_{i_2}, \ldots, X_{i_n}$ имеют то же совместное распределение, что и сл. в. $X_1, X_2, \ldots, X_n$. Поэтому все $n!$ событий $X_{i_1} < X_{i_2} < \ldots < X_{i_n}$ равновероятны. Более того, для любых промежутков I_{a_i, b_i}, $i = 1, \ldots, n$, равновероятны события

$$\{X_{i_1} < X_{i_2} < \ldots < X_{i_n},\ X_{i_k} - X_{i_{k-1}} \in I_{a_k, b_k},\ k = 1, \ldots, n\},\ X_{i_0} = 0.$$

Учитывая, что равенство $X_i = X_j$ при $i \neq j$ происходит с вероятностью 0, имеем для любых $y_i > 0$, $i = 1, \ldots, n$,

$$\mathbf{P}(\omega : (n - i + 1)(X_{(i)}(\omega) - X_{(i-1)}(\omega)) > y_i,\ i = 1, \ldots, n) =$$

$$= n!\mathbf{P}(\omega : X_1(\omega) < \ldots < X_n(\omega),\ (n - i + 1)(X_i(\omega) - X_{i-1}(\omega)) > y_i,\ i =$$

$$= 1, \ldots, n) = n!\mathbf{P}(\omega : X_i(\omega) - X_{i-1}(\omega) > y_i/(n - i + 1),\ i = 1, \ldots, n) =$$

$$= n! \int\limits_{\{x_i - x_{i-1} > y_i/(n-i+1),\, i = 1, \ldots, n\}} \!\!\!\cdots\!\!\! \int e^{-x_1 - \ldots - x_n}\, dx_1 \ldots dx_n =$$

$$= n! \int_{y_1/n}^{\infty} e^{-x_1}\, dx_1 \int_{x_1 + y_2/(n-1)}^{\infty} e^{-x_2}\, dx_2 \ldots \int_{x_{n-1} + y_n/1}^{\infty} e^{-x_n}\, dx_n = e^{-y_1} e^{-y_2} \ldots e^{-y_n}.]$$

4.18. З а д а ч а. Сл. в. X, Y имеют совместную плотность $f_{X,Y}(x, y)$. Показать, что сл. в. $Z=X+Y$ имеет плотность

$$f_Z(z)=\int_{-\infty}^{\infty} f_{X,Y}(x, z-x)\,dx.$$

Если X и Y независимы, то

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(x)\,f_Y(z-x)\,dx.$$

$$\Big[F_Z(z)=\iint\limits_{x+y\leqslant z} f_{X,Y}(x, y)\,dx\,dy=\int_{-\infty}^{\infty} dv\int_{-\infty}^{\infty} f_{X,Y}(u, v-u)\,du,$$

где сделана замена переменных $u=x$, $v=x+y$.]

4.19. З а д а ч а Сл. в. X_1, X_2 независимы и имеют стандартные гамма-распределения с параметрами p_1, p_2 соответственно:

$$f_{X_i}(x)=\Gamma(p_i)^{-1}x^{p_i-1}e^{-x},\quad x>0,\ i=1,\ 2.$$

Показать, что сл. в. $Y=X_1+X_2$ имеет гамма-распределение с параметром p_1+p_2.

$$\Big[f_Y(y)=\Gamma(p_1)^{-1}\,\Gamma(p_2)^{-1}\int_0^y x^{p_1-1}e^{-x}(y-x)^{p_2-1}e^{-(y-x)}\,dx.$$

Сделав замену переменной $x=yu$, получаем

$$f_Y(y)=ce^{-y}y^{p_1+p_2-1},\quad c=\Gamma(p_1)^{-1}\,\Gamma(p_2)^{-1}\int_0^1 u^{p_1-1}(1-u)^{p_2-1}\,du.$$

Так как интеграл от $f_Y(y)$ по всей прямой равен 1, то $c=$ $=\Gamma(p_1+p_2)^{-1}$. Отметим попутно доказанное равенство

$$\int_0^1 u^{p_1-1}(1-u)^{p_2-1}\,du=\Gamma(p_1)\,\Gamma(p_2)/\Gamma(p_1+p_2).\Big]$$

4.20. З а д а ч а. Сл. в. X_i, $i=1, \ldots, n$, независимы и имеют нормальное распределение с параметрами a_i, $\sigma_i>0$:

$$f_{X_i}(x)=\sigma_i^{-1}\varphi(\sigma_i^{-1}(x-a_i)),\quad \varphi(x)=\left(1/\sqrt{2\pi}\right)e^{-x^2/2}.$$

Показать, что сл. в. $Y=X_1+\ldots+X_n$ имеет нормальное распределение с параметрами $a=a_1+\ldots+a_n$, $\sigma=(\sigma_1^2+\ldots+\sigma_n^2)^{1/2}$, воспользовавшись 4.18 либо 4.15.

[Сл. в. $U_i=\sigma_i^{-1}(X_i-a_i)$, $i=1, \ldots, n$, независимы и имеют стандартные нормальные распределения, $Y=X_1+\ldots+X_n=\sigma_1U_1+a_1+$ $+\ldots+\sigma_nU_n+a_n=\sigma((\sigma_1/\sigma)U_1+\ldots+(\sigma_n/\sigma)U_n)+a\equiv\sigma Z+a$, так что достаточно проверить, что сл. в. Z является стандартной нормальной. Поскольку строка $(\sigma_1/\sigma, \ldots, \sigma_n/\sigma)$ имеет единичную евклидову норму, то ее можно дополнить до ортогональной матрицы и затем воспользоваться задачей 4.15.]

4.21. Задача. Сформулировать и решить задачи 4.1, 4.2 для случайных величин, обладающих плотностями.

[Если $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)=h_1(x_1)\ldots h_n(x_n)$, где h_i, $i=1, \ldots, n$, — некоторые неотрицательные функции, то сл. в. $X_1, \ldots, X_n$ независимы и т. д.]

Понятие независимости случайных величин было введено отдельно для дискретного и непрерывного случаев. Следующее определение объединяет оба варианта. Сл. в. $X_1, \ldots, X_n$ назовем *независимыми*, если для любых промежутков I_{a_i,b_i}, $-\infty \leqslant a_i \leqslant b_i \leqslant \infty$, $i=1, \ldots, n$

$$\mathbf{P}(\omega : X_i(\omega) \in I_{a_i,b_i},\ i=1, \ldots, n) = \prod_{i=1}^{n} P(\omega : X_i(\omega) \in I_{a_i,b_i}).$$

Наборы сл. в. $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_m$ назовем *независимыми между собой*, если

$$\mathbf{P}(\omega : X_i(\omega) \in I_{a_i,b_i},\ i=1, \ldots, n,\ Y_j(\omega) \in I_{c_j,d_j},\ j=1, \ldots, m) =$$
$$= \mathbf{P}(\omega : X_i(\omega) \in I_{a_i,b_i},\ i=1, \ldots, n)\,\mathbf{P}(\omega : Y_j(\omega) \in I_{c_j,d_j},\ j=1, \ldots, m)$$

для любых промежутков I_{a_i,b_i}, $i=1, \ldots, n$, I_{c_j,d_j}, $j=1, \ldots, m$.

4.22. Задача. Наборы непрерывных сл. в. $X_1, \ldots, X_n$ и $X_{n+1}, \ldots, X_{n+m}$ независимы между собой. Показать, что сл. в. $U=g(X_1, \ldots, X_n)$, $V=h(X_{n+1}, \ldots, X_{n+m})$ независимы.

[Полагая для краткости $n=m=1$, имеем

$$\mathbf{P}(\omega : U(\omega) \in I_{a,b},\ V(\omega) \in I_{c,d}) =$$
$$= \iint\limits_{\{(g(x_1),h(x_2)) \in I_{a,b}\times I_{c,d}\}} f_{X_1}(x_1) f_X(x_2)\, dx_1 dx_2 =$$
$$= \int\limits_{\{g(x_1) \in I_{a,b}\}} f_{X_1}(x_1)\, dx_1 \cdot \int\limits_{\{h(x_2) \in I_{c,d}\}} f_{X_2}(x_2)\, dx_2 =$$
$$= \mathbf{P}(\omega : U(\omega) \in I_{a,b})\,\mathbf{P}(\omega : V(\omega) \in I_{c,d}).]$$

4.23. Задача. Сл. в. $X_1, X_2, \ldots$ независимы и одинаково распределены со стандартной экспоненциальной плотностью $f(x)=e^{-x}$, $x>0$. Определим *пуассоновский поток событий* на полупрямой $t \geqslant 0$ (ср. 3.11—3.13), полагая $\tau_k=X_1+\ldots+X_k$, $k=1, 2, \ldots$, и

$$N(t)=n \Leftrightarrow \tau_n \leqslant t,\quad \tau_{n+1}>t,\quad n=0, 1, 2, \ldots,\ \tau_0=0.$$

Вычислить распределения вероятностей τ_k и $N(t)$.

[Используя 4.19 при $p_1=p_2=1$ и 4.22, индукцией по k получаем, что сл. в. τ_k имеет гамма-распределение с параметром k:

$$f_{\tau_k}(t)=(1/(k-1)!)\,t^{k-1}e^{-t},\quad t>0.$$

Ф. р. сл. в. τ_k равна

$$F_{\tau_k}(t)=1-e^{-t}(1+t/1!+\ldots+t^{n-1}/(n-1)!),$$

в чем легко убедиться дифференцированием. Далее,

$$\mathbf{P}(\omega: N(t;\,\omega)=n)=\mathbf{P}(\omega:\tau_n(\omega)\leqslant t,\ \tau_{n+1}(\omega)>t)=$$
$$=\mathbf{P}(\omega:\tau_n(\omega)\leqslant t)-\mathbf{P}(\omega:\tau_{n+1}(\omega)\leqslant t)=(t^n/n!)\,e^{-t}.]$$

4.24. Пример (времена ожидания в пуассоновском процессе). В условиях задачи 4.23 будем интерпретировать параметр t как временной и введем сл. в. $\tau_{N(t)+1}-t$ и $t-\tau_{N(t)}$, которые называют соответственно остающимся и прошедшим временами ожидания. При $N(t)=0$ на отрезке $[0,\,t]$ не было зарегистрировано событий случайного потока, и мы полагаем $\tau_0=0$, $t-\tau_0=t$. В остальных случаях величина $t-\tau_{N(t)}$ равняется времени, прошедшему с момента последнего перед t появления события пуассоновского потока.

Запишем совместную плотность сл. в. τ_n, $\tau_{n+1}=\tau_n+X_{n+1}$, учитывая, что пара τ_n, τ_{n+1} связана с парой τ_n, X_{n+1} линейным преобразованием с определителем 1, и применяя 3.26:

$$f_{\tau_n,\tau_{n+1}}(u,\,v)=f_{\tau_n,X_{n+1}}(u,\,v-u)=f_{\tau_n}(u)\cdot f_{X_{n+1}}(v-u)=$$
$$=(1/(n-1)!)\,u^{n-1}e^{-v},\quad 0\leqslant u\leqslant v,$$

где мы воспользовались независимостью τ_n и X_{n+1} и формулой для плотности $f_{\tau_n}(u)$, полученной в 4.23. Для вычисления совместного распределения сл. в. $t-\tau_{N(t)}$, $\tau_{N(t)+1}-t$ используем разложение:

$$\{\omega: t-\tau_{N(t)}\in I_{a,b},\ \tau_{N(t)+1}-t\in I_{c,d}\}=$$
$$=\bigcup_{n=0}^{\infty}\{\omega: t-\tau_{N(t)}\in I_{a,b},\ \tau_{N(t)+1}\in I_{c,d},\ N(t)=n\}=$$
$$=\bigcup_{n=0}^{\infty}\{\omega: t-\tau_n\in I_{a,b},\ \tau_{n+1}-t\in I_{c,d},\ N(t)=n\}.$$

При $a\geqslant 0$, $c>0$ событие $t-\tau_n\in I_{a,b}$, $\tau_{n+1}-t\in I_{c,d}$ влечет событие $\tau_n\leqslant t$, $\tau_{n+1}>t$, т. е. $N(t)=n$, откуда, продолжая равенство, имеем

$$\bigcup_{n=0}^{\infty}\{\omega: t-\tau_n\in I_{a,b},\ \tau_{n+1}-t\in I_{c,d}\}.$$

Напомним, что в используемой нами модели бесконечной последовательности независимых случайных величин вероятностная мера задана лишь на таких подмножествах элементарных событий, которые описывают поведение конечного числа членов последовательности. Подмножество $\{\omega : t-\tau_{N(t)} \in I_{a,b},\ \tau_{N(t)+1}-t \in I_{c,d}\}$, очевидно, не входит в указанный класс, зато оно представлено в виде счетного объединения попарно непересекающихся событий из этого класса. Кажется вполне логичным определить вероятность рассматриваемого события как сумму ряда из вероятностей событий в упомянутом представлении:

$$\sum_{n=0}^{\infty} \mathbf{P}(\omega : t-\tau_n \in I_{a,b},\ \tau_{n+1}-t \in I_{c,d}) =$$

$$= \sum_{n=0}^{\infty} \mathbf{P}(\omega : \tau_n \in I_{t-b,t-a},\ \tau_{n+1} \in I_{t+c,t+d}).$$

При $n \geqslant 1$, $0 \leqslant a < b \leqslant t$, $0 \leqslant c < d$ для члена ряда имеем выражение

$$\int_{t-b}^{t-a} \int_{t+c}^{t+d} (1/(n-1)!)\, u^{n-1} e^{-v} du dv.$$

Сумма этих выражений по n от 1 до ∞ равна

$$\int_{t-b}^{t-a} \int_{t+c}^{t+d} e^{u} e^{-v}\, du dv = \int_{a}^{b} e^{-x} dx \int_{c}^{d} e^{-y} dy.$$

При $n=0$ получаем

$$\mathbf{P}(\omega : t-\tau_0 = t,\ \tau_1 - t \in I_{c,d}) = \mathbf{P}(\omega : \tau_1 \in I_{t+c,t+d}) =$$
$$= (1-e^{-t-d}) - (1-e^{-t-c}) = e^{-t}(e^{-c}-e^{-d}).$$

Таким образом,

$$\mathbf{P}(\omega : t-\tau_{N(t)} \in I_{a,b},\ \tau_{N(t)+1}-t \in I_{c,d}) = Q_1^t(I_{a,b})\, Q_2(I_{c,d}),$$

где Q_2 — мера с плотностью e^{-x}, $x>0$, $Q_1{}^t$ — вероятностная мера смешанного типа: она имеет атом величины e^{-t} в точке t, т. е. вероятностная масса e^{-t} приписана точке t, а остальная вероятностная масса распределена непрерывно на отрезке $[0, t]$ с плотностью e^{-x}. Полагая по очереди $a=0$, $b=\infty$ и $c=0$, $d=\infty$, получаем

$$\mathbf{P}(\omega : t-\tau_{N(t)} \in I_{a,b}) = Q_1^t(I_{a,b}),\quad \mathbf{P}(\omega : \tau_{N(t)+1}-t \in I_{c,d}) = Q_2(I_{c,d}),$$

так что прошедшее и остающееся времена ожидания являются независимыми случайными величинами, причем последнее экспоненциально распределено. Когда $t\to\infty$, атом e^{-t} исчезает, и прошедшее время ожидания в пределе оказывается экспоненциальным. Именно с этой ситуацией мы имели дело в примере 3.11, где предельный случай достигался за счет рассмотрения пуассоновского процесса, начавшегося на $-\infty$.

4.25. Задача. В обозначениях задачи 4.23 показать, что сл. в. $\eta_k=\tau_k/\tau_{n+1}$, $k=1, \ldots, n$, распределены так же, как порядковые статистики для n независимых равномерно распределенных на отрезке $[0, 1]$ случайных величин (см. 3.24).

[Пользуясь формулой 3.26 преобразования плотности при линейном преобразовании $\tau_k=X_1+\ldots+X_k$, $k=1, 2, \ldots, n+1$, случайных величин (ср. 4.24), получаем

$$f_{\tau_1,\tau_2,\ldots,\tau_{n+1}}(t_1, t_2, \ldots, t_{n+1})=e^{-t_{n+1}}, \quad 0<t_1<t_2<\ldots<t_{n+1}.$$

Полагая для краткости $n=2$ и выбирая $0<a_1<b_1\leqslant a_2<b_2<1$, запишем

$$\mathbf{P}(\omega : \eta_1(\omega)\in I_{a_1,b_1},\ \eta_2(\omega)\in I_{a_2,b_2})\equiv\iiint\limits_{\{A\}} e^{-t_3}\,dt_1dt_2dt_3,$$

где интегрирование ведется по области

$$A=\{t_1/t_3\in I_{a_1,b_1},\ t_2/t_3\in I_{a_2,b_2},\ t_1<t_2<t_3\}=$$
$$=\{t_1/t_3\in I_{a_1,b_1},\ t_2/t_3\in I_{a_2,b_2}\}.$$

Заменяя тройной интеграл повторным, получаем

$$\int\limits_0^1 e^{-t_3}dt_3\int\limits_{a_2t_3}^{b_2t_3}dt_2\int\limits_{a_1t_3}^{b_1t_3}dt_1=\Gamma(3)(b_1-a_1)(b_2-a_2).$$

Отсюда, как и в 3.24, заключаем, что $f_{\eta_1,\eta_2}(t_1, t_2)=2!$, $0<t_1<t_2<1$.]

4.26. Пример (пуассоновский процесс). Пуассоновский поток событий (см. 4.23) однозначно представляется семейством сл. в. $N(t;\omega)$, $t\geqslant 0$. При каждом $t>0$ ($N(t;\omega)$ равно числу зарегистрированных событий на отрезке $[0, t]$ ($N(0;\omega)=0$), при фиксированном ω $N(t;\omega)$ есть неубывающая ступенчатая функция переменного t с единичными скачками (рис. 21). *Случайный процесс* $N(t;\omega)$, $t\geqslant 0$, называется *пуассоновским*. В 4.23 показано, что

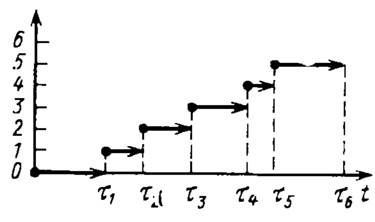

Рис. 21

$$\mathbf{P}(\omega : N(t;\omega)=n)=(t^n/n!)e^{-t}, \quad t>0, \quad n=0, 1, 2, \ldots.$$

Совместные распределения вероятностей значений процесса в конечном числе точек t_i, $i=1, 2, \ldots, k$,

$$\mathbf{P}(\omega : N(t_i;\omega)=n_i,\ i=1, 2, \ldots, k)$$

называют его *конечномерными распределениями*. Система всех конечномерных распределений случайного процесса выполняет ту же роль, что и совместное распределение вероятностей конечного набора случайных величин: она определяет вероятности любых событий, связанных с поведением процесса, и с ее помощью можно задать непосредственно процесс с данными конечномерными распределениями.

Покажем, что пуассоновский процесс имеет *независимые приращения*: для любых k, $0=t_0<t_1<\ldots<t_k$ сл. в. $N(t_i)-N(t_{i-1})$, $i=1, 2, \ldots, k$, независимы, и отсюда выведем его конечномерные распределения. В свою очередь независимость приращений пуассоновского процесса легко следует из следующего свойства пуассоновского потока событий. Возьмем произвольную точку $t>0$ и рассмотрим части пуассоновского потока на промежутках $[0, t]$ и $[t, \infty)$. Оказывается, что эти части независимы, а поток, высекаемый промежутком $[t, \infty)$, — пуассоновский с началом в точке t.

Итак, пусть $X_1, X_2, \ldots$ — последовательность независимых случайных величин с одинаковой экспоненциальной плотностью распределения,

$$\tau_k=X_1+\ldots+X_k,\ k=1, 2, \ldots, \tau_0=0,$$

$$\{N(t)=n\}\Leftrightarrow\{\tau_n\leqslant t,\ \tau_{n+1}>t\},$$

где $t>0$ и $N(0)=0$. Положив $Y_1=\tau_{N(t)+1}-t$, $Y_i=\tau_{N(t)+i}-\tau_{N(t)+i-1}$, $i=2, 3, \ldots$, получаем, что события пуассоновского потока после момента t регистрируются в моменты $t+Y_1$, $t+Y_1+Y_2, \ldots$, так что требуется показать, что сл. в. $Y_1, Y_2, \ldots$ независимы, экспоненциально распределены и не зависят от потока на отрезке $[0, t]$, т. е. от сл. в. $\tau_1, \ldots, \tau_{N(t)}$, $N(t)$. Соответствующие рассуждения близки к проведенным в примере 4.24.

Независимость частей пуассоновского потока, высекаемых промежутками $[0, t]$ и $[t, \infty)$, сводится к соотношениям

$$\mathbf{P}(\omega : N(t)=n,\ \tau_i\in I_{a_i,b_i},\ i\leqslant n,\ Y_k\in I_{c_k,d_k},\ k\leqslant m)=$$
$$=\mathbf{P}(\omega : N(t)=n,\ \tau_i\in I_{a_i,b_i},\ i\leqslant n)\,\mathbf{P}(\omega : Y_k\in I_{c_k,d_k},\ k\leqslant m)$$

при любых $n\geqslant 0$, $m\geqslant 1$ и промежутках $I\subseteq[0, \infty)$. При $n>0$ перепишем левую часть соотношения с учетом независимости сл. в. $X_1, X_2, \ldots$ в виде произведения

$$\mathbf{P}(\omega : (\tau_1, \ldots \tau_{n+1})\in A)\cdot\mathbf{P}(\omega : X_{n+k}\in I_{c_k,d_k},\ 2\leqslant k\leqslant m)\equiv p_1p_2,$$
$$A=\{(t_1, \ldots, t_{n+1}) : t_i\in I_{a_i,b_i},\ i\leqslant n,\ t_n<t,\ t_{n+1}-t\in I_{c_1,d_1}\}.$$

Используя выражение для совместной плотности сл. в. $\tau_1, \ldots, \tau_{n+1}$ из 4.25 и предполагая, что $0\leqslant a_1<b_1\leqslant a_2<b_2\leqslant\ldots\leqslant a_n<b_n\leqslant\leqslant t$, получаем

$$p_1=\int\limits_A\ldots\int e^{-t_{n+1}}\,dt_1\ldots dt_{n+1}=\int\limits_{a_1}^{b_1}\ldots\int\limits_{a_n}^{b_n}dt_1\ldots dt_n\int\limits_{c_1+t}^{d_1+t}e^{-t_{n+1}}dt_{n+1},$$

$$p_1p_2=\int\limits_{a_1}^{b_1}\dots\int\limits_{a_n}^{b_n}e^{-t}dt_1\dots dt_n\cdot\int\limits_{c_1}^{d_1}\dots\int\limits_{c_m}^{d_m}e^{-u_1-\dots-u_m}du_1\dots du_m.$$

Полагая $c_i = 0$, $d_i = \infty$, $i = 1, \dots, m$, находим отсюда, что (при $b_i \leqslant a_{i+1}$)

$$\mathbf{P}(\omega : N(t)=n,\ \tau_i \in I_{a_i,b_i},\ i=1,\dots,n)=\int\limits_{a_1}^{b_1}\dots\int\limits_{a_n}^{b_n}e^{-t}dt_1\dots dt_n.$$

Вводя функцию $h(t_1, \dots, t_n) = e^{-t}$ при $0 \leqslant t_1 < \dots < t_n \leqslant t$ и $h(t_1, \dots, t_n) = 0$ при остальных значениях $t_1, \dots, t_n$, можно записать для множеств $B \subseteq R^n$

$$\mathbf{P}(\omega : N(t)=n,\ (\tau_1,\dots,\tau_n)\in B)=\int\dots\int\limits_{\{B\}}h(t_1,\dots,t_n)\,dt_1\dots dt_n,$$

$$\mathbf{P}(\omega : N(t)=n,\ (\tau_1,\dots,\tau_n)\in B,\ Y_k\in I_{c_k,d_k},\ k\leqslant m)=$$

$$=\mathbf{P}(\omega : N(t)=n,\ (\tau_1,\dots,\tau_n)\in B)\cdot\int\limits_{c_1}^{d_1}\dots\int\limits_{c_m}^{d_m}e^{-u_1-\dots-u_m}du_1\dots du_m.$$

Полагая $B = R^n$ и суммируя по n, получаем

$$\sum_{n=1}^{\infty}\mathbf{P}(\omega : N(t)=n,\ Y_k\in I_{c_k,d_k},\ k\leqslant m)=$$

$$=(1-e^{-t})\int\limits_{c_1}^{d_1}\dots\int\limits_{c_m}^{d_m}e^{-u_1-\dots-u_m}du_1\dots du_m.$$

При $n = 0$ имеем

$$\mathbf{P}(\omega : N(t)=0,\ Y_k\in I_{c_k,d_k},\ k\leqslant m)=$$

$$=\mathbf{P}(\omega : X_1\geqslant t,\ X_1-t\in I_{c_1,d_1},\ X_k\in I_{c_k,d_k},\ 2\leqslant k\leqslant m)=$$

$$=e^{-t}\int\limits_{c_1}^{d_1}\dots\int\limits_{c_m}^{d_m}e^{-u_1-\dots-u_m}du_1\dots du_m.$$

Событие $\{Y_k\ \ I_{c_k,d_k},\ k\leqslant m\}$ представим в виде счетного объединения попарно непересекающихся событий $\{N(t)=n,\ Y_k\in I_{c_k,d_k},\ k\leqslant m\}$, $n = 0, 1, 2, \dots$. Как и в 4.24, *определим* вероятность этого события как сумму ряда

$$\sum_{n=0}^{\infty} \mathbf{P}(\omega : N(t)=n,\ Y_k \in I_{c_k, d_k},\ k \leqslant m)=$$

$$=\int_{c_1}^{d_1} \ldots \int_{c_m}^{d_m} e^{-u_1-\ldots-u_m} du_1 \ldots du_m,$$

что заканчивает доказательство независимости двух частей пуассоновского потока, высекаемых промежутками $[0, t]$ и $[t, \infty)$, и устанавливает пуассоновость части потока на $[t, \infty)$.

Из доказанного вытекает, что при любом $s>t$ сл. в. $N(t)$, $N(s)-N(t)$ независимы и $N(s)-N(t)$ имеет пуассоновское распределение с параметром $s-t$. В общем случае при $0<t_1<t_2<\ldots<t_k$, применяя последовательно утверждение о независимости частей пуассоновского потока на $[0, t_1]$ и $[t_1, \infty)$, $[t_1, t_2]$ и $[t_2, \infty)$ и т. д, заключаем, что сл. в. $N(t_i)-N(t_{i-1})$, $i=1, \ldots, k$, независимы и имеют пуассоновские распределения с параметрами t_i-t_{i-1}, $i=1, \ldots, k$. Отсюда получаем конечномерные распределения пуассоновского процесса:

$$\mathbf{P}(\omega : N(t_i)=n_i,\ i \leqslant k)=\mathbf{P}(\omega : N(t_i)-N(t_{i-1})=n_i-n_{i-1},\ i \leqslant k)=$$

$$=\prod_{i=1}^{k} ((t_i-t_{i-1})^{n_i-n_{i-1}}/(n_i-n_{i-1})!)\, e^{-(t_i-t_{i-1})},\ n_1 \leqslant \ldots \leqslant n_k.]$$

4.27. Пример (случайное блуждание и броуновское движение). Пусть $X_1, X_2, \ldots$ — независимые случайные величины, $\mathbf{P}(\omega : X_i(\omega)=1)=p$, $0<p<1$, $\mathbf{P}(\omega : X_i(\omega)=-1)=1-p$. *Простое случайное блуждание,* начинающееся из нуля, определяется как последовательность случайных величин

$$S_0=0, \quad S_n=X_1+\ldots+X_n, \quad n=1, 2, \ldots .$$

Блуждание с началом в произвольной целой точке a есть последовательность $a+S_n(\omega)$, $n=0, 1, 2, \ldots$. Случайное блуждание $S_n(\omega)$, $n=0, 1, 2, \ldots$, представляет собой *случайный процесс с дискретным параметром (временем)*. Так как при любых

$$0=n_0<n_1<n_2<\ldots<n_k \text{ сл. в. } S_{n_i}-S_{n_{i-1}}=X_{n_{i-1}+1}+\ldots+X_{n_i},$$
$$i=1, \ldots, k,$$

независимы, то этот процесс имеет независимые приращения. При $p=1/2$ блуждание называется симметричным. Оно было использовано в § 3 при эвристическом определении процесса броуновского движения W_t.

В 3.19 было установлено для случайного процесса с непрерывным временем

$$W_t^{(h)}=S_{[nt]}/\sqrt{n}=(X_1+\ldots+X_{[nt]})/\sqrt{n}, \quad h=1/\sqrt{n}, \quad t \geqslant 0,$$

где $[a]$ — целая часть числа a, $p=1/2$, что

$$\mathbf{P}(\omega : W_t^{(h)} \leqslant x) \to \int_{-\infty}^{x} (1/\sqrt{t})\,\varphi(y/\sqrt{t})\,dy, \quad n \to \infty,$$

$\varphi(x)$ — стандартная нормальная плотность. Это привело к предположению, что случайный процесс, описывающий одномерную проекцию броуновского движения, должен удовлетворять условию: плотность распределения вероятностей значения процесса в момент времени t равна

$$f_{W_t}(x) = (1/\sqrt{t})\,\varphi(x/\sqrt{t}).$$

При любых $0 = t_0 < t_1 < \ldots < t_k$ сл. в.

$$W_{t_i}^{(h)} - W_{t_{i-1}}^{(h)} = (X_{[nt_{i-1}]+1} + \ldots + X_{[nt_i]})/\sqrt{n}, \quad i = 1, \ldots, k,$$

независимы, так что

$$\mathbf{P}(\omega : W_{t_i}^{(h)} - W_{t_{i-1}}^{(h)} \leqslant x_i,\ i \leqslant k) = \prod_{i=1}^{k} \mathbf{P}(\omega : W_{t_i}^{(h)} - W_{t_{i-1}}^{(h)} \leqslant x_i) \to$$

$$\to \prod_{i=1}^{k} \int_{-\infty}^{x_i} (1/\sqrt{t_i - t_{i-1}})\,\varphi(y_i/\sqrt{t_i - t_{i-1}})\,dy_i.$$

Таким образом, от процесса W_t нужно еще потребовать независимости приращений, причем приращение $W_t - W_s$, $s < t$, должно быть нормальным с плотностью $(1/\sqrt{t-s})\,\varphi(x/\sqrt{t-s})$. Положив $\eta_i = W_{t_i} - W_{t_{i-1}}$, имеем

$$f_{\eta_1,\ldots,\eta_k}(y_1, \ldots, y_k) = \prod_{i=1}^{k} (1/\sqrt{t_i - t_{i-1}})\,\varphi(y_i/\sqrt{t_i - t_{i-1}}).$$

Линейное преобразование $W_{t_i} = \eta_1 + \ldots + \eta_i$, $i = 1, \ldots, k$, имеет определитель, равный единице, откуда, применяя формулу преобразования плотности 3.26, получаем

$$f_{W(t_1),\ldots,W(t_k)}(x_1, \ldots, x_k) = f_{\eta_1,\ldots,\eta_k}(x_1, x_2 - x_1, \ldots, x_k - x_{k-1}) =$$

$$= \prod_{i=1}^{k} (1/\sqrt{t_i - t_{i-1}})\,\varphi((x_i - x_{i-1})/\sqrt{t_i - t_{i-1}}), \quad x_0 = t_0 = 0.$$

Итак, дискретное приближение процесса броуновского движения приводит к предположению, что процесс $W_t(\omega)$ должен иметь указанные конечномерные распределения (с точностью до масштабного преобразования). Остается открытым вопрос о построении вероятностного пространства, на котором могло бы быть определено семейство сл. в. $W_t(\omega)$, $t \geqslant 0$, с заданными конечномерными распределениями, о чем будет речь ниже (см. 10.41).

§ 5.

УСЛОВНАЯ ВЕРОЯТНОСТЬ

Имеется немало практических ситуаций, когда наблюдатель располагает информацией о значениях, которые приняли некоторые «наблюдаемые» сл. в. $Y(\omega)$, $Z(\omega)$, ..., а интерес представляет некоторая недоступная непосредственному наблюдению сл. в. $X(\omega)$. Пусть, например, известно, что $Y(\omega)\in I_{a,b}$, $Z(\omega)\in I_{c,d}$, ..., т. е. произошло событие $B=\{\omega : Y(\omega)\in I_{a,b},\ Z(\omega)\in I_{c,d}, \ldots\}$. Возникает вопрос, как изменятся наши представления о случайном объекте $X(\omega)$, если пространство элементарных событий сокращено до некоторого своего подмножества B? Так, в схеме повторного выбора без возвращения из урны с M красными, $N-M$ черными шарами допустим, что первый вытащенный шар красный. Тогда при втором извлечении красный шар появляется с вероятностью $(M-1)/(N-1)$, а черный — с вероятностью $(N-M)/(N-1)$. Если же цвет первого извлеченного шара неизвестен, то возможности появления красного и черного шара при втором извлечении характеризуются вероятностями M/N и $(N-M)/N$ (см. 1.5). При выборе с возвращением информация о результатах предшествующих извлечений никак не отражается на возможностях следующего извлечения, и вероятности появления красного и черного шара остаются равными M/N и $(N-M)/N$ соответственно.

Предположим сначала, что пространство элементарных событий Ω соответствует опыту с N равновозможными исходами. Пусть известно, что произошло некоторое событие B. Кажется правдоподобным, что сокращение пространства Ω до множества B не должно приводить к изменению относительных возможностей появлений различных элементарных событий $\omega\in B$. Это позволяет считать все M_B элементарных событий, составляющих B, равновозможными. Тем самым приходим к новому пространству элементарных событий B и определенной на подмножествах $A\subseteq B$ вероятностной мере

$$\mathbf{P}_B(A)=M_A/M_B.$$

Поделив числитель и знаменатель в правой части на N, перепишем формулу, определяющую $\mathbf{P}_B(A)$, в виде

$$\mathbf{P}_B(A)=\mathbf{P}(A)/\mathbf{P}(B), \quad A\subseteq B,$$

где $\mathbf{P}$ — вероятностная мера на исходном пространстве элементарных событий Ω. Вероятностную меру $\mathbf{P}_B$ полезно распространить на все подмножества Ω, полагая $\mathbf{P}_B(\omega)=0$ при $\omega\in\overline{B}$, что приводит к формуле

$$\mathbf{P}_B(A)=\mathbf{P}(AB)/\mathbf{P}(B), \quad A\subseteq\Omega.$$

При любых фиксированных A и $B\neq\varnothing$ число $\mathbf{P}_B(A)$ называют *условной вероятностью* события A при условии события B, вероятностная мера $\mathbf{P}_B$ называется *условной* (при условии события B). Для условной вероятности наряду с $\mathbf{P}_B(A)$ используют обозначение $\mathbf{P}(A|B)$.

Условная вероятность, как и обычная, безусловная, допускает частотную интерпретацию. Если некоторый опыт повторяется многократно, то закон устойчивости частот утверждает, что частоты ν_A, ν_B, ν_{AB} событий A, B, AB приближаются к вероятностям:

$$\nu_A \approx \mathbf{P}(A), \ \nu_B \approx \mathbf{P}(B), \quad \nu_{AB} \approx \mathbf{P}(AB).$$

Выделив лишь те испытания, которые привели к событию B, для частоты события A в этих испытаниях получаем выражение

$$\nu_{AB}/\nu_B \approx \mathbf{P}(AB)/\mathbf{P}(B) = \mathbf{P}(A|B).$$

Таким образом, частота появления события A в подпоследовательности испытаний, при которых наблюдается B, приближается с ростом числа испытаний к условной вероятности $\mathbf{P}(A|B)$.

Отмеченные соображения сохраняют свое значение и для произвольного дискретного вероятностного пространства. Если произошло событие B, то относительные возможности осуществления любого $\omega \in B$ по-прежнему характеризуются распределением вероятностей $p(\omega)$, суженным на B, но удобнее нормировать $p(\omega)$, $\omega \in B$, перейдя к $p_B(\omega) = p(\omega)/\mathbf{P}(B)$. Функция $p_B(\omega)$ удовлетворяет условию нормировки $\Sigma_{\omega \in B} p_B(\omega) = 1$ и $\mathbf{P}_B(A) = \mathbf{P}(A)/\mathbf{P}(B)$, $A \subseteq B$, — вероятностная мера на B. Определим $p_B(\omega)$ на всем Ω, полагая $p_B(\omega) = 0$ при $\omega \in \overline{B}$, что соответствует невозможности осуществления $\omega \in \overline{B}$, когда произошло B. В результате приходим к *определению условной вероятности*

$$\mathbf{P}_B(A) \equiv \mathbf{P}(A|B) = \mathbf{P}(AB)/\mathbf{P}(B), \quad A, B \subseteq \Omega, \quad \mathbf{P}(B) > 0,$$

имеющему ту же форму, что и для опытов с равновозможными исходами. В такой форме определение условной вероятности сохраняется для произвольного вероятностного пространства.

Итак, наряду с основным вероятностным пространством $(\Omega, p(\omega))$ имеем для любого $B \subseteq \Omega$, $\mathbf{P}(B) > 0$ вероятностное пространство $(\Omega, p_B(\omega))$. Распределение вероятностей сл. в. $X(\omega)$, рассматриваемой на $(\Omega, p_B(\omega))$, называется условным при условии события B. Взяв $B_y = \{\omega : Y(\omega) = y\}$, получаем условное распределение сл. в. X при условии $Y = y$:

$$p_{X|Y}(x|y) = \mathbf{P}(\{\omega : X(\omega) = x\} | \{\omega : Y(\omega) = y\}) = p_{X,Y}(x, y)/p_Y(y).$$

При любом y, для которого $p_Y(y) > 0$, $p_{X|Y}(x|y)$ представляет собой дискретное распределение вероятностей по x. Аналогично определяется условное распределение набора сл. в. $\mathbf{X} = (X_1, \ldots, X_n)$ относительно набора сл. в. $\mathbf{Y} = (Y_1, \ldots, Y_m)$.

5.1. Задача. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — последовательность испытаний Бернулли с вероятностью успеха p, $T = X_1 + \ldots + X_n$. Найти условное распределение $\mathbf{X}$ относительно T.

[При $t=x_1+\ldots+x_n$ имеем

$$p_{\mathbf{x},T}(\mathbf{x}, t)/(p_T(t)=p_{\mathbf{x}}(\mathbf{x})/p_T(t)=p^t(1-p)^{n-t}/(C_n^t p^t(1-p)^{(n-t)}),$$

т. е. распределение $p_{\mathbf{x}|T}(\mathbf{x}|t)$ сосредоточено на множестве C_n^t наборов $\mathbf{x}=(x_1, \ldots, x_n)$ из 0 и 1 с t единицами и является равномерным на нем.]

5.2. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и одинаково распределены по Пуассону:

$$\mathbf{P}_{X_i}(k)=(\lambda^k/k!)e^{-\lambda}, \quad k=0, 1, 2, \ldots.$$

Найти условное распределение набора сл. в. $\mathbf{X}=(X_1, \ldots, X_n)$ относительно сл. в. $T=X_1+\ldots+X_n$.

[Используя 4.7, имеем при $t=x_1+\ldots+x_n$, $\mathbf{x}=(x_1, \ldots, x_n)$

$$p_{\mathbf{X}|T}(\mathbf{x}|t)=p_{\mathbf{X}}(\mathbf{x})/p_T(t)=(\lambda^{x_1+\ldots+x_n}/(x_1! \ldots x_n!))\times$$
$$\times e^{-n\lambda}(((n\lambda)^t/t!)e^{-n\lambda})^{-1}=(t!/(x_1! \ldots x_n!))n^{-t}.$$

Таким образом, при условии события $T=t$ сл. в. $X_1, \ldots, X_n$ имеют такое же распределение вероятностей, как если бы X_i представляло число шаров в i-м ящике, $i=1,\ldots,n$, при случайном размещении по ним t шаров (см. 1.25).]

5.3. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и одинаково равномерно распределены с множеством значений $1, 2, \ldots, N$ (повторный случайный выбор с возвращением). Найти условное распределение $\mathbf{X}=(X_1, \ldots, X_n)$ при условии $X_{(n)}=\max(X_1, \ldots, X_n)$.

$$[\mathbf{P}(\omega : X_{(n)}(\omega)=k)=\mathbf{P}(\omega : X_i(\omega)\leqslant k,\ i\leqslant n)-\mathbf{P}(\omega : X_i(\omega)\leqslant k-1,\ i\leqslant n)=$$
$$=(k/N)^n-((k-1)/N)^n,\ p_{\mathbf{X}|X_{(n)}}(\mathbf{x}|k)=(k^n-(k-1)^n)^{-1},$$
$$1\leqslant k\leqslant N, \quad \mathbf{x}=(x_1, \ldots, x_n), \max(x_1, \ldots, x_n)=k.]$$

В математической статистике изучают методы извлечения информации из результатов наблюдений $x_1, x_2, \ldots, x_n$ за сл. в. $X_1, X_2, \ldots, X_n$, чтобы сделать выводы о неизвестных параметрах вероятностной модели, описывающей $X_1, X_2, \ldots, X_n$. Точнее говоря, предполагается, что на пространстве $\mathscr{X}=\{\mathbf{x}\}$ возможных наблюдений $\mathbf{x}=(x_1, x_2, \ldots, x_n)$ задано семейство вероятностных распределений $\mathbf{P}_\theta$, зависящее от некоторого параметра θ, и что наблюдение $\mathbf{x}$ получено в результате вероятностного эксперимента, порождающего меру $\mathbf{P}_\theta$ на $\mathscr{X}$ с некоторым неизвестным наблюдателю значением θ.

Понятно, что наличие в *выборке* $\mathbf{x}$ информации о неизвестном истинном значении θ связано с зависимостью распределения вероятностей на $\mathscr{X}$ от θ. Если бы на самом деле эксперимент не зависел от θ, то нечего было бы рассчитывать на оценивание θ по $\mathbf{x}$. Отсюда возникает идея «разложить» эксперимент на части, содержащие и не содержащие информацию о θ. В задачах 5.1—5.3 это удается сделать. Так, в 5.1 распределение вероятностей сл. в. $T=X_1+\ldots+X_n$ — биномиальное с параметрами n, p, а условное распределение $X_1, \ldots, X_n$ при условии T не зависит от параметра p. Это означает, что если известно, что $T=t$, то уточнение значений, которые при этом приняли сл. в. $X_1, \ldots, X_n$, не вносит дополнительной информации о параметре p. Поэтому при решении задачи оценивания неизвестного p по наблюдениям $\mathbf{x}=(x_1, \ldots, x_n)$ можно без ущерба перейти от $\mathbf{x}$ к $t=x_1+\ldots+x_n$. Дальнейший анализ упрощается за счет того, что от точки $\mathbf{x}$ n-мерного пространства мы перешли к точке t на прямой. Параметр p является вероятностью в схеме Бернулли, и закон устойчивости частот служит основанием для оценивания p с помощью $(x_1+\ldots+x_n)/n$.

Проведенные выше рассуждения показывают, что во всяком случае этот способ оценивания использует всю полезную для оценивания информацию, содержащуюся в выборке. В задаче 5.3 аналогичную роль выполняет сл. в. $X_{(n)}=\max(X_1, \ldots, X_n)$. В урновой интерпретации речь идет об оценке неизвестного числа шаров в урне по результатам n-кратного случайного выбора с возвращением (шары занумерованы числами от 1 до N). Как видно, для оценки N достаточно знать $x_{(n)}=\max(x_1, \ldots, x_n)$. В статистике функцию от результатов наблюдений, содержащую всю полезную информацию о параметре модели, называют *достаточной статистикой*. Достаточность статистики $h(\mathbf{x})$ означает, что условное распределение выборки $\mathbf{x}=(x_1, \ldots, x_n)$ при условии $h(\mathbf{x})=y$ одно и то же для любого распределения вероятностей $\mathbf{P}_\theta$ из заданного семейства.

Определение условной вероятности можно записать в виде

$$\mathbf{P}(AB)=\mathbf{P}(A|B)\,\mathbf{P}(B),$$

или для распределений вероятностей случайных величин

$$p_{X,Y}(x, y)=p_{Y|X}(y|x)\,p_X(x).$$

В такой форме это определение можно использовать для построения модели зависимых опытов. Представим себе, что на некотором множестве $\mathscr{X}$ задано распределение вероятностей $p_1(x)$, а для каждого $x\in\mathscr{X}$ на множестве $\mathscr{Y}$ задано распределение вероятностей $p_2(y|x)$. Определим на множестве $\mathscr{X}\times\mathscr{Y}$ функцию $p(x, y)=p_2(y|x)\,p_1(x)$. Так как

$$\sum_{\{(x,y)\in\mathscr{X}\times\mathscr{Y}\}} p(x, y)=\sum_{\{x\in\mathscr{X}\}} p_1(x)\sum_{\{y\in\mathscr{Y}\}} p_2(y|x)=1,$$

то $p(x, y)$ — распределение вероятностей на $\mathscr{X}\times\mathscr{Y}$. Если $p_2(y|x)$ не является постоянной по x (на подмножестве $x: p_1(x)>0$), то $p(x, y)$ не распадается в произведение $p_1(x)p_2(y)$, и вероятностная модель $(\mathscr{X}\times\mathscr{Y}, p(x, y))$ отлична от модели независимых испытаний.

Используем изложенные соображения, чтобы дать дополнительное пояснение понятия достаточной статистики. Возьмем для примера модель испытаний Бернулли. Пусть $\mathcal{T}=\{0, 1, \ldots, n\}$, $({}_1pt)=C_n^t p^t(1-p)^{n-t}$, $t\in\mathcal{T}$, $\mathcal{X}=\{\mathbf{x}\}$, $\mathbf{x}=(x_1, \ldots, x_n)$, $x_i=0$ или 1, $i=1, \ldots, n$, $p_2(\mathbf{x}|t)=0$, если $x_1+\ldots+x_n\neq t$, и $p_2(\mathbf{x}|t)=1/C_n{}^t$ при $x_1+\ldots+x_n=t$. На множестве $\mathcal{T}\times\mathcal{X}$ определим $p(t, \mathbf{x})=$ $=p_1(t)p_2(\mathbf{x}|t)$. Распределение вероятностей $p(t, \mathbf{x})$ сосредоточено на парах $(t, \mathbf{x})$, таких, что $t=x_1+\ldots+x_n$, и вероятность такой пары равна $p^{x_1+\ldots+x_n}(1-p)^{n-x_1-\ldots-x_n}$. Легко понять, что модель $(\mathcal{T}\times\mathcal{X}, p(t, \mathbf{x}))$ эквивалентна модели испытаний Бернулли. При этом опыт $(\mathcal{T}, p_2(t))$ зависит от параметра p — вероятности успеха, в то время как «условный» опыт $(\mathcal{X}, p_2(\mathbf{x}|t))$, приводящий к значению $\mathbf{x}$ пары $(t, \mathbf{x})$, от параметра p не зависит. Следовательно, для определения, каково истинное p, элементом $\mathbf{x}$ пары $(t, \mathbf{x})$ можно пренебречь.

5.4. Задача. Сл. в. X, Y независимы и имеют биномиальные распределения с параметрами n, p и m, p соответственно. Найти условное распределение X при условии $X+Y$.

[Так как сл. в. $X+Y$ имеет биномиальное распределение с параметрами $n+m$, p (см. 4.5), то

$$P_{X|X+Y}(k|l)=P(\omega : X(\omega)=k,\quad X(\omega)+Y(\omega)=l)\times$$
$$\times(P(\omega : X(\omega)+Y(\omega)=l))^{-1}=C_n^k p^k(1-p)^{n-k}C_m^{l-k}p^{l-k}(1-p)^{m-l+k}\times$$
$$\times(C_{n+m}^l p^l(1-p)^{n+m-l})^{-1}=C_n^k C_m^{l-k}/C_{n+m}^l.]$$

5.5. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и имеют одинаковые геометрические распределения вероятностей $p_{X_i}(k)=$ $=(1-p)^k p$, $k=0, 1, 2, \ldots$. Найти условное распределение $\mathbf{X}=$ $=(X_1, \ldots, X_n)$ при условии $T=X_1+\ldots+X_n$.

[Ввиду 4.6 $p_T(t)=C_{n+t-1}^t p^n(1-p)^t$, откуда при $x_1+\ldots+x_n=t$, $p_{\mathbf{X}|T}(\mathbf{x}|t)=p_{\mathbf{X}}(\mathbf{x})/p_T(t)=1/C_{n+t-1}^t$.]

5.6. Задача. Рассмотрим пример выборочного контроля продукции, предполагая, что каждое изделие независимо от остальных является стандартным с вероятностью p_1 и, проходя контроль, независимо от остальных проверяется с вероятностью p_2. Обозначим U число изделий, прошедших стол контроля (как проверенных, так и непроверенных), прежде чем обнаружено первое бракованное изделие, и V количество бракованных (и непроверенных) изделий из числа этих U изделий (считается, что при проверке брак обнаруживается). Найти условное распределение V при условии U.

[Опишем рассматриваемый опыт двумя независимыми последовательностями испытаний Бернулли $X_1, X_2, \ldots, Y_1, Y_2, \ldots$, где $X_i=1$, если i-е изделие — стандартное, и $X_i=0$, если оно бракованное; $Y_i=1$, если i-е изделие подвергается контролю, и $Y_i=0$ — в противном случае, так что $\mathbf{P}(\omega : X_i(\omega)=1)=p_1$, $\mathbf{P}(\omega : Y_i(\omega)=1=p_2$.

Отсюда

$$p_U(n)=\mathbf{P}(\omega:(X_{n+1}(\omega),\ Y_{n+1}(\omega))=(0,\ 1),\ (X_i(\omega),\ Y_i(\omega))\neq(0,\ 1),\ i\leqslant n)=$$
$$=(1-p_1)\,p_2\,(1-(1-p_1)\,p_2)^n,$$

$$p_{U,V}(n,\ k)=\mathbf{P}(\omega:(X_{n+1},\ Y_{n+1})=(0,\ 1),\ (X_1,\ Y_1,\ \ldots,\ X_n,\ Y_n)\in B_k),$$

где B_k состоит из всех последовательностей $(x_1,\ y_1,\ \ldots,\ x_n,\ y_n)$, что равно k пар $(x_i,\ y_i)=(0,\ 0)$, а остальные пары $(x_i,\ y_i)=(1,\ 0)$ или $(1,\ 1)$. Таким образом,

$$p_{U,V}(n,\ k)=(1-p_1)\,p_2\,C_n^k\,((1-p_1)(1-p_2))^k\,(p_1(1-p_2)+p_1p_2)^{n-k}.$$

Полагая $\mathbf{P}(\{\omega:(X_1(\omega),\ Y_1(\omega))=(0,\ 0)\}\,|\,\{\omega:(X_1(\omega),\ Y_1(\omega))\neq 0,\ 1)\})=$ $=(1-p_1)(1-p_2)/(1-(1-p_1)\,p_2)=1-p$, получаем, что условное распределение

$$p_{V|U}(k\,|\,n)=C_n^k(1-p)^k\,p^{n-k},\quad k=0,\ 1,\ \ldots,\ n,$$

— биномиальное.]

5.7. Задача. Пусть некоторое насекомое откладывает случайное число яиц с пуассоновским распределением вероятностей. Предположим, что из отложенных яиц насекомые развиваются независимо и вероятность развития насекомого из яйца равна p (с вероятностью $1-p$ развитие не происходит). Найти распределение вероятностей числа новых насекомых.

[Результат опыта можно описать парой целых неотрицательных чисел $(k,\ l)$, где k — число отложенных яиц, l — число развившихся насекомых. Положим $\Omega=\{(k,\ l)\}$, $X(\omega)=k$, $Y(\omega)=l$, $\omega=(k,\ l)$. Условие задачи можно выразить в виде следующих соотношений:

$$p_X(k)=(\lambda^k/k!)\,e^{-\lambda},\quad p_{Y|X}(l\,|\,k)=C_k^l p^l(1-p)^{k-l},\quad l,\ k=0,\ 1,\ 2,\ \ldots.$$

Это позволяет построить вероятностную модель опыта, положив

$$p(k,\ l)=p_{Y|X}(k,\ l)\,p_X(k)=((p\lambda)^l/l!)\,e^{-\lambda}\,(\lambda(1-p))^{k-l}/(k-l)!$$

Совместное распределение сл. в. $X(\omega)$, $Y(\omega)$, определенных на вероятностном пространстве $(\Omega,\ p(\omega))$, равно $p_{X,Y}(k,\ l)=p(k,\ l)$, откуда

$$p_Y(l)=\sum_{k=1}^{\infty}p_{X,Y}(k,\ l)=((p\lambda)^l/l!)\,e^{-\lambda}\sum_{k=l}^{\infty}(\lambda(1-p))^{k-l}/(k-l)!=$$
$$=((p\lambda)^l/l!)\,e^{-\lambda}\,e^{\lambda(1-p)}=((p\lambda)^l/l!)\,e^{-p\lambda},\quad l=0,\ 1,\ 2,\ \ldots.]$$

5.8. Задача. Пусть вероятность p_n, что в случайно выбранной семье n детей, равна $p_n=\alpha p^n$, $n=1,\ 2,\ \ldots,\ p_0=1-(p_1+p_2+\ldots)$. Предположим, что в семье из n детей все 2^n комбинаций полов равновероятны. Найти распределение вероятностей числа мальчиков в случайно выбранной семье.

[Представим результат опыта строчкой $\omega=(n, x_1, \ldots, x_n)$, $n=0$, $1, 2, \ldots$, $x_i=0$ или 1 $i=1, \ldots n$ ($\omega=(0)$ при $n=0$), полагая n равным числу детей и $x_i=1$, если i-й ребенок в семье мальчик, $x_i=0$, если девочка. Пусть $N(\omega)=n$ при $\omega=(n, x_1, \ldots, x_n)$ и запишем условие задачи в виде

$$p_N(n)=p_n, \quad p(\omega\,|\,N=n)=2^{-n}, \quad \omega=(n, x_1, \ldots, x_n).$$

На пространстве элементарных событий $\Omega=\{\omega\}$ введем распределение вероятностей

$$p(\omega)=p(\omega\,|\,N=n)\,p_N(n)=2^{-n}p_n, \quad \omega=(n, x_1, \ldots, x_n).$$

Пусть $S(\omega)=x_1+\ldots+x_n$, $\omega=(n, x_1, \ldots, x_n)$, $S(\omega)=0$, $\omega=(0)$ — число мальчиков в семье. Тогда

$$P_{N,S}(n, k)=\sum_{x_1+\ldots+x_n=k} p(\omega)=C_n^k 2^{-n}p_n, \quad n\geqslant 1, \quad p_{N,S}(0, 0)=p_0,$$

$$p_S(0)=p_0+\sum_{n=1}^{\infty} 2^{-n}\alpha p^n=p_0+\alpha p/(2-p),$$

$$p_S(k)=\sum_{n=k}^{\infty} p_{N,S}(n, k)=\sum_{n=k}^{\infty} C_n^k 2^{-n}\alpha p^n, \quad k\geqslant 1.$$

Используя 4.9, имеем

$$p_S(k)=\alpha(p/2)^k(1-p/2)^{-k-1}=2\alpha p^k(2-p)^{-k-1}, \quad k\geqslant 1.]$$

Определяя понятие условной вероятности, мы исходили из представления, что если с данным опытом связаны сл. в. $X(\omega)$, $Y(\omega)$ и $X(\omega)$ наблюдаема, то из наблюдения за сл. в. X, вообще говоря, можно извлечь информацию о ненаблюдаемой сл. в. Y. Источник этой информации – связь между функциями $X(\omega)$ и $Y(\omega)$ через общее значение ω. Если известно, что $X(\omega)=x$, то при этом сокращается множество возможных исходов опыта ω, а следовательно, изменяются, вообще говоря, возможности появления тех или иных значений сл. в. $Y(\omega)$. При данном x условное распределение $p_{Y|X}(y\,|\,x)$ в сравнении с безусловным $p_Y(y)$ отражает наличие информации в X относительно Y. В частном случае, когда $p_{Y|X}(y\,|\,x)$ от x не зависит: $p_{Y|X}(y\,|\,x)=h(y)$, где $h(y)$ – некоторая функция переменного y (при таких x, что $p_X(x)>0$), из соотношения

$$p_{X,Y}(x, y)=p_{Y|X}(y\,|\,x)p_X(x)=p_X(x)\,h(y)$$

и задачи 4.1 выводим, что $h(y)=p_Y(y)$ и

$$p_{X,Y}(x, y)=p_X(x)\,p_Y(y).$$

Таким образом, из соотношения

$$p_{Y|X}(y|x)=p_Y(y)$$

вытекает, что сл. в. X и Y независимы и наоборот. То же утверждение сохраняется для наборов сл. в. $\mathbf{X}=(X_1, \ldots, X_n)$, $\mathbf{Y}=(Y_1, \ldots, Y_m)$:

$$p_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})=p_{\mathbf{X}}(\mathbf{x})p_{\mathbf{Y}}(\mathbf{y}) \Leftrightarrow p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})=p_{\mathbf{Y}}(\mathbf{y}),$$

де $\mathbf{x}=(x_1, \ldots, x_n)$, $\mathbf{y}=(y_1 \ldots, y_n)$.

5.9. Задача. Доказать, что $\mathbf{X}=(X_1, \ldots, X_n)$ и $\mathbf{Y}=(Y_1, \ldots, Y_m)$ независимы между собой тогда и только тогда, когда для любых функций $h(\mathbf{x})$ и $g(\mathbf{y})$ независимы сл. в. $h(\mathbf{X})$ и $g(\mathbf{Y})$.

[Необходимость устанавливается аналогично 4.3. Обратно, предположив, что для некоторых $\mathbf{x}^\circ$, $\mathbf{y}^\circ$, $p_{\mathbf{x},\mathbf{y}}(\mathbf{x}^\circ, \mathbf{y}^\circ)\neq p_{\mathbf{x}}(\mathbf{x}^\circ)p_{\mathbf{x}}(\mathbf{y}^\circ)$, определим $h(\mathbf{x})=1$ при $\mathbf{x}=\mathbf{x}^\circ$, $g(\mathbf{y})=1$ при $\mathbf{y}=\mathbf{y}^\circ$ и положим $h(\mathbf{x})=0$, $g(\mathbf{y})=0$ для остальных значений $\mathbf{x}$, $\mathbf{y}$. Тогда

$$\mathbf{P}(\omega : h(X(\omega))=1,\ g(Y(\omega))=1)=p_{\mathbf{X},\mathbf{Y}}(\mathbf{x}^0, \mathbf{y}^0)\neq$$

$$\neq p_{\mathbf{X}}(x^0)p_{\mathbf{Y}}(y^0)=\mathbf{P}(\omega : h(\mathbf{X}(\omega))=1)\mathbf{P}(\omega : g(\mathbf{Y}(\omega))=1),$$

т. е. сл. в. $h(\mathbf{X})$ и $g(\mathbf{Y})$ — зависимые.]

5.10. Задача. Доказать, что сл. в. $X_1, X_2, \ldots, X_n$ независимы, если при любом k сл. в. X_k не зависит от набора сл. в. $(X_1, \ldots, X_{k-1})$, $k=2, \ldots, n$.

[По условию $p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)=p_{X_1,\ldots,X_{n-1}}(x_1, \ldots, x_{n-1})p_{X_n}(x_n)=$

$$=p_{X_1,\ldots,X_{n-2}}(x_1, \ldots, x_{n-2})p_{X_{n-1}}(x_{n-1})p_{X_n}(x_n)=\ldots=$$

$$=p_{X_1}(x_1)\ldots p_{X_{n-1}}(x_{n-1})p_{X_n}(x_n).]$$

Совместное распределение произвольного набора сл. в. $X_1, \ldots, X_n$ всегда можно представить в форме

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)=p_{X_1}(x_1)p_{X_2|X_1}(x_2|x_1)\times$$

$$\times p_{X_3|X_1,X_2}(x_3|x_1, x_2)\ldots p_{X_n|X_1,\ldots,X_{n-1}}(x_n|x_1, \ldots, x_{n-1}),$$

где рассматриваются лишь такие наборы $x_1, \ldots, x_n$ значений сл. в. $X_1, \ldots, X_n$, которые имеют положительную вероятность: $p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)>0$ (а следовательно $p_{X_1,\ldots,X_k}(x_1, \ldots, x_k)>0$ при всех $k=1, \ldots, n$).

Кроме важного частного случая, когда все условные вероятности $p_{X_{k+1}|X_1,\ldots,X_k}(x_{k+1}|x_1, \ldots, x_k)$ не зависят от $x_1, \ldots, x_k$ и, следовательно, сл. в. $X_1, \ldots, X_n$ независимы, имеется другой важный случай, когда условные вероятности $p_{X_{k+1}|X_1,\ldots,X_k}(x_{k+1}|x_1, \ldots$ $\ldots x_k)$ зависят от $x_1, \ldots, x_{k-1}$ лишь формально, т. е. представляют собой некоторые функции $p_k(x_k, x_{k+1})$ только переменных x_k, x_{k+1}. В этом случае говорят, что сл. в. $X_1, \ldots, X_n$ связаны *марковской зависимостью,* или сл. в. $X_1, \ldots, X_n$ образуют *марковскую цепь,* по имени русского ученого А. А. Маркова, впервые исследовавшего такой характер зависимости с общих позиций. Формула, выражающая совместное распределение через произведение условных, для марковских цепей приобретает вид

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = p_{X_1}(x_1)\, p_1(x_1, x_2) \ldots p_{n-1}(x_{n-1}, x_n).$$

Верно и обратное утверждение: если

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = p_1(x_1)\, p_2(x_1, x_2) \ldots p_{n-1}(x_{n-1}, x_n),$$

где p_i являются распределениями вероятностей по x_i, $i=1, \ldots, n$, то последовательность $X_1, \ldots, X_n$ образует марковскую цепь. Действительно, суммируя совместное распределение сл. в. $X_1, \ldots$ $\ldots, X_n$ по всем значениям x_n, находим

$$p_{X_1,\ldots,X_{n-1}}(x_1, \ldots, x_{n-1}) = p_1(x_1)\, p_2(x_1, x_2) \ldots p_{n-1}(x_{n-2}, x_{n-1}),$$

откуда

$$p_{X_n\, X_1,\ldots,X_{n-1}}(x_n \mid x_1, \ldots, x_{n-1}) = p_n(x_{n-1}, x_n)$$

и т. д.; в частности, $p_{X_1}(x_1) = p_1(x_1)$.

5.11. Пример (модель Эренфестов процесса диффузии). Представим себе, что урна, содержащая N шаров, мысленно разделена на две половины. Каждую единицу времени случайно выбирается один из N шаров и переносится из своей половины в другую. При n-кратном повторении опыта имеется N^n равновероятных элементарных событий, описывающих выбор шаров. Введем сл. в. X_k — число шаров в какой-нибудь одной половине урны после k-го извлечения, $k=1, 2, \ldots, n$. Подчеркнем, что X_k — функции элементарного события в схеме повторного выбора с возвращением, если указаны номера шаров, лежащих в выделенной половине урны в начальный момент времени. Пусть число шаров в этой половине вначале равно $X_0=a$.

Подсчитаем число элементарных событий, удовлетворяющих условию $X_1=x_1, \ldots, X_n=x_n$. Это число ненулевое, если все $x_1, \ldots$ $\ldots, x_n$ заключены между 0 и N, а соседние члены последовательности отличаются ровно на единицу. Рассматривая наборы, удовлетворяющие указанным требованиям, заметим, что на очередном шаге переходу от $1 \leqslant x_{i-1} \leqslant N-1$ к $x_i=x_{i-1}+1$ соответствует $N-x_{i-1}$ возможностей выбора шара из урны (выбирается шар с номером, не содержащимся среди номеров шаров в выделенной половине), переходу к $x_i=x_{i-1}-1$ отвечает x_{i-1} возможностей, переходам от $x_{i-1}=0$ к $x_i=1$ и от $x_{i-1}=N$ к $x_i=N-1$ отвечает любая из N возможностей.

Введем набор чисел $p_{k,l}$, k, $l=0, 1, \ldots, N$, полагая $p_{k,k+1}=1-k/N$, $p_{k,k-1}=k/N$ при $k=0, 1, \ldots, N$ и $p_{k,l}=0$ при остальных k и l. Легко видеть, что число элементарных событий, составляющих событие $\{\omega : X_i(\omega)=x_i, i=1, \ldots, n\}$ при начальном числе шаров в выделенной половине a, равно $Np_{a,x_1}Np_{x_1,x_2} \ldots Np_{x_{n-1},x_n}$, откуда получаем совместное распределение сл. в. $X_1, \ldots, X_n$ в виде

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = p_{a,x_1}p_{x_1,x_2} \ldots p_{x_{n-1},x_n},$$

из которого следует, что эти величины образуют цепь Маркова.

Тот факт, что условное распределение

$$p_{X_k|X_1,\ldots,X_{k-1}}(x_k|x_1, \ldots, x_{k-1}) = p_{x_{k-1},x_k}$$

зависит лишь от сл. в. X_{k-1}, означает, что информация о X_k, содержащаяся в сл. в. $X_1, \ldots, X_{k-1}$, сосредоточена в сл. в. X_{k-1}. Это легко понять и без всяких вычислений. Условие $X_{k-1}=x_{k-1}$ означает, что в выделенной половине содержится x_{k-1} шаров. Поэтому имеется x_{k-1} возможностей из N получить $X_k=x_{k-1}-1$ и $N-x_{k-1}$ возможностей из N получить $X_k=x_{k-1}+1$, все это независимо от предыстории процесса до момента k.

Покажем, как можно моделировать марковскую цепь $X_1, \ldots, X_n$, исходя непосредственно из чисел $p_{k,l}$, через которые выражается их совместное распределение. Зададим следующий механизм блуждания по точкам $0, 1, \ldots, N$. Свяжем с каждым $k=0, 1, \ldots, N$ урну с N шарами, из которых $N-k$ шаров помечены числом $+1$ и k шаров — числом -1 (при $k=0$ и $k=N$ все шары помечены одним числом). Предположим, что если блуждающая частица находится в положении j, то производится однократный выбор с возвращением из урны с номером j, и частица смещается на единицу в направлении, отмеченном на извлеченном шаре. При $j=0$ и $j=N$ направление определяется однозначно, так что урны с этими номерами введены для единообразия. Выбрав некоторое a в качестве начального положения, обозначим $\hat{X}_k$ положение частицы в момент времени k. Легко видеть, что из N^n элементарных событий опыта длительности n имеется $N^n p_{a,x_1}p_{x_1,x_2} \ldots p_{x_{n-1},x_n}$ элементарных событий, приводящих к траектории $(a, x_1, x_2, \ldots, x_n)$, так что

$$p_{\hat{X}_1,\hat{X}_2,\ldots,\hat{X}_n}(x_1, x_2, \ldots, x_n) = p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n).$$

В построенной модели зависимость сл. в. $\hat{X}_k$ от предыстории процесса $\hat{X}_1, \ldots, \hat{X}_{k-1}$ только через последнюю сл. в. $\hat{X}_{k-1}$ проглядывается особенно отчетливо. Ввиду вероятностной эквивалентности наборов сл. в. $(\hat{X}_1, \ldots, \hat{X}_n)$ и $(X_1, \ldots, X_n)$ эти соображения переносятся и на исходный процесс.

5.12. Задача. Пусть $X_0, X_1, \ldots, X_n$ — марковская цепь. Показать, что при $k=0, 1, \ldots, n$ имеет место следующая формула:

$$p_{X_0,X_1,\ldots,X_k}(x_0, x_1, \ldots, x_k) = p_{X_0}(x_0)\, p_{X_1|X_0}(x_1|x_0) \ldots p_{X_k|X_{k-1}}(x_k|x_{k-1}).$$

$$[p_{X_0,\ldots,X_k}(x_0, \ldots, x_k) =$$

$$= p_{X_k|X_0,\ldots,X_{k-1}}(x_k|x_0, \ldots, x_{k-1})\, p_{X_0,\ldots,X_{k-1}}(x_0, \ldots, x_{k-1}).$$

Суммируя по всем возможным значениям $x_0, x_1, \ldots, x_{k-2}$ и учитывая, что $p_{X_k|X_0,\ldots,X_{k-1}}(x_k|x_0, \ldots, x_{k-1})$ зависит от $x_0, x_1, \ldots$ $\ldots, x_{k-2}$ лишь формально, получаем для совместного распределения сл. в. X_{k-1}, X_k:

$$p_{X_{k-1},X_k}(x_{k-1}, x_k) = p_{X_k|X_0,\ldots,X_{k-1}}(x_k|x_0, \ldots, x_{k-1})\, p_{X_{k-1}}(x_{k-1}),$$

откуда

$$p_{X_k|X_0,X_1,\ldots,X_{k-1}}(x_k|x_0, x_1, \ldots, x_{k-1}) = p_{X_k|X_{k-1}}(x_k|x_{k-1}).$$

Индукция по k завершает доказательство.]

Часто используемую формулу

$$\mathbf{P}(A) = \sum_{\{k\}} \mathbf{P}(AB_k),$$

где B_k, $k=1, 2, \ldots$, — полная группа событий, полезно переписать в виде

$$\mathbf{P}(A) = \sum_{\{k\}} \mathbf{P}(A|B_k)\,\mathbf{P}(B_k).$$

Это так называемая *формула полной вероятности*. Для дискретных сл. в. X, Y (или наборов сл. в. $\mathbf{X}=(X_1, \ldots, X_n)$, $\mathbf{Y}=(Y_1, \ldots, Y_m)$) имеем отсюда

$$p_X(x) = \sum_{\{y\}} p_{X|Y}(x|y)\, p_Y(y),$$

где суммирование ведется по всем значениям y сл. в. Y.

5.13. Задача. Вывести из 5.12 индукцией по k, что

$$p_{X_k,\ldots,X_n}(x_k, \ldots, x_n) = p_{X_k}(x_k)\, p_{X_{k+1}|X_k}(x_{k+1}|x_k) \ldots p_{X_n|X_{n-1}}(x_n|x_{n-1}).$$

[Выполним первый шаг индукции, суммируя соотношение 5.12 по всем значениям x_0:

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \left(\sum_{\{x_0\}} p_{X_0}(x_0)\, p_{X_1|X_0}(x_1|x_0)\right) p_{X_2|X_1}(x_2|x_1) \times$$

$$\times \ldots p_{X_n|X_{n-1}}(x_n|x_{n-1}) = p_{X_1}(x_1)\, p_{X_2|X_1}(x_2|x_1) \ldots p_{X_n|X_{n-1}}(x_n|x_{n-1})].$$

5.14. Задача. Показать, что для марковской цепи выполняется соотношение

$$p_{X_{k+1},\ldots,X_n|X_0,\ldots,X_k}(x_{k+1},\ldots,x_n|x_0,\ldots,x_k)=$$
$$=p_{X_{k+1},\ldots,X_n|X_k}(x_{k+1},\ldots,x_n|x_k).$$

[Из 5.12, 5.13 получаем

$$p_{X_{k+1},\ldots,X_n|X_0,\ldots,X_k}(x_{k+1},\ldots,x_n|x_0,\ldots,x_k)=p_{X_0,\ldots,X_n}(x_0,\ldots,x_n)\times$$
$$\times(p_{X_0,\ldots,X_k}(x_0,\ldots,x_k))^{-1}=p_{X_{k+1}|X_k}(x_{k+1}|x_k)\ldots p_{X_n|X_{n-1}}(x_n|x_{n-1})=$$
$$=(p_{X_k}(x_k))^{-1}p_{X_k,\ldots,X_n}(x_k,\ldots,x_n)=p_{X_{k+1},\ldots,X_n|X_k}(x_{k+1},\ldots,x_n|x_k).]$$

5.15. Задача. Показать, что для марковской цепи «будущее» $\mathbf{X}=(X_{k+1},\ldots,X_n)$ и «прошлое» $\mathbf{Y}=(X_0,\ldots,X_{k-1})$ *условно независимы* при заданном «настоящем» X_k, т. е.

$$p_{\mathbf{X},\mathbf{Y}|X_k}(\mathbf{x},\mathbf{y}|x_k)=p_{\mathbf{X}|X_k}(\mathbf{x}|x_k)\,p_{\mathbf{Y}|X_k}(\mathbf{y}|x_k),$$

где $\mathbf{x}=(x_0,\ldots,x_{k-1})$, $\mathbf{y}=(x_{k+1},\ldots,x_n)$.

[Используя 5.14, получаем

$$p_{\mathbf{X},\mathbf{Y}|X_k}(\mathbf{x},\mathbf{y}|x_k)=p_{\mathbf{X},\mathbf{Y},X_k}(\mathbf{x},\mathbf{y},x_k)(p_{X_k}(x_k))^{-1}=$$
$$=p_{\mathbf{X}|X_k,\mathbf{Y}}(\mathbf{x}|x_k,\mathbf{y})\cdot p_{X_k,\mathbf{Y}}(x_k,\mathbf{y})(p_{X_k}(x_k))^{-1}=p_{\mathbf{X}|X_k}(\mathbf{x}|x_k)p_{\mathbf{Y}|X_k}(\mathbf{y}|x_k).]$$

Последовательность сл. в. $X_0, X_1, \ldots, X_n$ была названа марковской цепью, если при любом k

$$p_{X_{k+1}|X_0,\ldots,X_k}(x_{k+1}|x_0,\ldots,x_k)=p_k(x_k,x_{k+1}),$$

где $p_k(x_k,x_{k+1})$ есть функция только переменных x_k, x_{k+1}. В 5.12 установлено, что $p_k(x_k,x_{k+1})=p_{X_{k+1}|X_k}(x_{k+1}|x_k)$, , так что определение марковской цепи можно дать в форме

$$p_{X_{k+1}|X_0,\ldots,X_k}(x_{k+1}|x_0,\ldots,x_k)=p_{X_{k+1}|X_k}(x_{k+1}|x_k).$$

Особенный интерес представляют марковские модели, для которых функции p_k при всех k одинаковы:

$$p_k(x,y)=p_{X_{k+1}|X_k}(y|x)=p(x,y)\equiv p_{x,y}\equiv p(y|x).$$

Вероятности $p(x,y)\equiv p_{x,y}$ называют *переходными вероятностями* марковской цепи. Подчеркнем, что при каждом фиксированном x $p(x,y)$ представляет собой распределение вероятностей по y.

В примере 5.11 $p_{x,x+1}=1-x/N$, $p_{x,x-1}=x/N$, $x=0, 1, \ldots, N$, остальные $p_{x,y}=0$. Простое случайное блуждание $S_n=a+X_1+\ldots$ $\ldots+X_n$, $n=0, 1, 2, \ldots$, где $X_1, X_2, \ldots$ независимы,

$p_{X_k}(1)=p$, $p_{X_k}(-1)=1-p$, $k=1, 2, \ldots$, является марковской цепью:

$$\mathbf{P}_{S_{k+1}|S_0,\ldots,S_k}(s_{k+1}|s_0, \ldots, s_k)=\mathbf{P}(\omega : S_i(\omega)=s_i,\ 0\leqslant i\leqslant k+1)\times$$
$$\times(\mathbf{P}(\omega : S_i(\omega)=s_i,\ 0\leqslant i\leqslant k))^{-1}=\mathbf{P}(\omega : S_i(\omega)=s_i,\ 0\leqslant i\leqslant k,$$
$$X_{k+1}(\omega)=s_{k+1}-s_k)\cdot(\mathbf{P}(\omega : S_i(\omega)=s_i,\ 0\leqslant i\leqslant k))^{-1}=$$
$$=\mathbf{P}(\omega : X_{k+1}(\omega)=s_{k+1}-s_k),$$

где на последнем шаге была использована независимость сл. в. X_{k+1} от набора сл. в. $S_0, \ldots, S_k$. Переходные вероятности этой марковской цепи равны $p_{i,i+1}=p$, $p_{i,i-1}=1-p$, остальные $p_{ij}=0$.

5.16. Задача. Пусть N белых и N черных шаров разложены по двум урнам так, что в обеих урнах число шаров одинаковое. Производится повторный случайный выбор из обеих урн, при котором вытащенные шары обмениваются местами. Обозначим X_k, $k=1, 2, \ldots$, число белых шаров в первой урне после k-го извлечения, и пусть $X_0=a$ — начальное число белых шаров в этой урне. Показать, что $X_0, X_1, \ldots, X_n$ — марковская цепь, выписать ее переходные вероятности.

[Опыт длительности n имеет N^{2n} равновероятных элементарных событий. Событие $\{\omega : X_i(\omega)=x_i,\ i=1, \ldots, n\}$ имеет ненулевую вероятность, только при $|x_{i+1}-x_i|\leqslant 1$, $i=0, 1, \ldots, n-1$, $x_0=a$. На i-м шаге переходу от $1\leqslant x_i\leqslant N-1$ к x_i+1 соответствует выбор черного шара из первой урны ($N-x_i$ возможностей) и белого шара из второй урны ($N-x_i$ возможностей), что дает $(N-x_i)^2$ из N^2 возможностей. Аналогично переход к $x_{i+1}=x_i-1$ происходит в x_i^2 случаях, а к $x_{i+1}=x_i$ — в $2x_i(N-x_i)$ случаях. При $x_i=0$ и $x_i=N$ величина x_{i+1} может принимать лишь два значения, для которых можно воспользоваться уже выведенными формулами для числа возможностей. Положим

$$p_{j,j+1}=(1-j/N)^2,\quad p_{j,j-1}=(j/N)^2,\quad p_{j,j}=2(j/N)(1-j/N)$$

при $j=0, 1, \ldots, N$ и $p_{kl}=0$ — при остальных k, l. Число элементарных событий, составляющих событие $\{\omega : X_i(\omega)=x_i,\ i=1, \ldots, n\}$, можно записать в виде $N^{2n}p_{a,x_1}p_{x_1,x_2}\ldots p_{x_{n-1},x_n}$, откуда

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)=p_{a,x_1}p_{x_1x_2}\ldots p_{x_{n-1},x_n},$$

так что $X_1, \ldots, X_n$ — марковская цепь с переходными вероятностями p_{kl}.]

Для непрерывной вероятностной модели определение условной вероятности $\mathbf{P}(A|B)$, как уже отмечалось, имеет такую же форму, что и для дискретной. Что касается непрерывных случайных величин, то, как известно, роль распределения вероятностей в дискретном случае здесь выполняет плотность распределения вероятностей. Соответственно вместо условных распределений рассматриваются *условные плотности*.

Пусть X, Y — пара случайных величин (или наборов сл. в. $\mathbf{X}=(X_1, \ldots, X_n)$, $\mathbf{Y}=(Y_1, \ldots, Y_m)$), $f_{X,Y}(x, y)$ — их совместная плотность, определенная на множестве $\mathscr{X}\times\mathscr{Y}$, так что

$$\iint\limits_{\mathscr{X}\times\mathscr{Y}} f_{X,Y}(x, y)\,dx\,dy=1.$$

Предположим дополнительно, что $f_Y(y)>0$ при всех $y\in\mathscr{Y}$ (это требование является излишним в общей теории, а в конкретных случаях его нетрудно удовлетворить, выделив некоторое подмножество $\mathscr{Y}'\subseteq\mathscr{Y}$, на котором $f_Y(y)>0$ и $\mathbf{P}(\omega : Y(\omega)\in\mathscr{Y}')=1$). *Условной плотностью* X при условии Y называется плотность распределения на $\mathscr{X}$, определяемая формулой

$$f_{X|Y}(x\,|\,y)=f_{X,Y}(x, y)/f_Y(y)$$

и зависящая от $y\in\mathscr{Y}$ как от параметра. Так как

$$\int\limits_{\mathscr{X}} f_{X|Y}(x\,|\,y)\,dx=\int\limits_{\mathscr{X}} f_{X,Y}(x, y)\,dx/f_Y(y)=1,$$

то $f_{X|Y}(x|y)$ действительно является плотностью вероятностного распределения по x. Отметим следующее соотношение:

$$f_X(x)=\int\limits_{\mathscr{Y}} f_{X,Y}(x, y)\,dy=\int\limits_{\mathscr{Y}} f_{X|Y}(x\,|\,y)\,f_Y(y)\,dy,$$

которое представляет собой непрерывный аналог формулы полной вероятности.

5.17. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и равномерно распределены на отрезке $[0, \theta]$, $X_{(1)}\leqslant X_{(2)}\leqslant \ldots \leqslant X_{(n)}$ — порядковые статистики (см. 4.17): $X_{(k)}$ равна k-му по величине в порядке возрастания среди значений $X_1, \ldots, X_n$. Найти условную плотность набора сл. в. $X_{(1)}, \ldots, X_{(n-1)}$ при условии сл. в. $X_{(n)}$.

[Из 3.24 следует, что

$$f_{X_{(1)},\ldots, X_{(n)}}(x_1, \ldots, x_n)=n!\theta^{-n},\ 0\leqslant x_1\leqslant \ldots \leqslant x_n\leqslant\theta,$$

и равна нулю при остальных значениях аргументов. Так как

$$\mathbf{P}(\omega : X_{(n)}(\omega)\leqslant t)=\mathbf{P}(\omega : X_i(\omega)\leqslant t,\ i=1, \ldots, n)=(t/\theta)^n,\ 0\leqslant t\leqslant\theta,$$

то при $0\leqslant x_1\leqslant \ldots \leqslant x_n\leqslant\theta$

$$f_{X_{(n)}}(x_n)=nx_n^{n-1}\theta^{-n},\ f_{X_{(1)},\ldots,X_{(n-1)}|X_{(n)}}(x_1, \ldots, x_{n-1}\,|\,x_n)=(n-1)!/x_n^{n-1}.$$

Заметим, что условная плотность сл. в. $X_{(1)}, \ldots, X_{(n-1)}$ при условии $X_{(n)}=x_n$ совпадает с плотностью $n-1$ порядковых статистик для $n-1$ равномерно распределенных на $[0, x_n]$ независимых случайных величин.]

5.18. Задача. Сл. в. $X_1, \ldots, X_{n+1}$ независимы и имеют стандартную экспоненциальную плотность, $\tau_k=X_1+\ldots+X_k$, $k=1, \ldots$

..., $n+1$. Найти условную плотность набора сл. в. $\tau_1, \ldots, \tau_n$ при условии сл. в. τ_{n+1}.

[Из 4.23 и 4.25 имеем

$$f_{\tau_{n+1}}(t_{n+1}) = (1/n!)\, t_{n+1}^n e^{-t_{n+1}}, \quad t_{n+1} > 0,$$

$$f_{\tau_1, \ldots, \tau_{n+1}}(t_1, \ldots, t_{n+1}) = e^{-t_{n+1}}, \quad 0 < t_1 < \ldots < t_{n+1},$$

$$f_{\tau_1, \ldots, \tau_n | \tau_{n+1}}(t_1, \ldots, t_n | t_{n+1}) = n!\, t_{n+1}^{-n}, \quad 0 < t_1 < \ldots < t_{n+1}.]$$

Аналогично дискретному случаю многомерную плотность можно разложить в произведение условных плотностей

$$f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) f_{X_2|X_1}(x_2 | x_1) \times$$
$$\times f_{X_3|X_1, X_2}(x_3 | x_1, x_2) \ldots f_{X_n|X_1, \ldots, X_{n-1}}(x_n | x_1, \ldots, x_{n-1}).$$

Если плотности $f_{X_{k+1}|X_1, \ldots, X_k}(x_{k+1} | x_1, \ldots, x_k)$ как функции $x_1, \ldots, x_k$ постоянны, $k=1, \ldots, n-1$, то сл. в. $X_1, \ldots, X_n$ являются независимыми. Как и в дискретном случае, *марковская зависимость* определяется требованием, чтобы плотности $f_{X_{k+1}|X_1, \ldots, X_k}(x_{k+1} | x_1, \ldots, x_k)$ зависели от $x_1, \ldots, x_{k-1}$ лишь формально, $k=2, \ldots, n-1$. При этом говорят, что сл. в. $X_1, \ldots, X_n$ образуют *непрерывную марковскую цепь*. Свойства дискретных марковских цепей, выведенные в 5.12—5.16, остаются верными и для непрерывных цепей с заменой условных распределений на условные плотности. Доказательства сохраняются с заменой, где это нужно, сумм на интегралы. Отметим, в частности, равенство

$$f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) f_{X_2|X_1}(x_2 | x_1) \ldots f_{X_n|X_{n-1}}(x_n | x_{n-1}),$$

которое может быть принято за определение марковской цепи; $f_{X_{k+1}|X_k}(x_{k+1} | x_k)$ называют *переходными плотностями*.

В качестве примера непрерывной цепи можно предложить последовательность $W_{t_1}, W_{t_2}, \ldots, W_{t_n}$, $t_1 < t_2 < \ldots < t_n$, значений процесса броуновского движения. Как следует из 4.27, переходные плотности этой цепи равны

$$f_{W(t_{i+1})|W(t_i)}(x_{i+1} | x_i) = (t_{i+1} - t_i)^{-1/2} \varphi\left((x_{i+1} - x_i)/\sqrt{t_{i+1} - t_i}\right).$$

5.19. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и имеют плотности $f_{X_i}(x_i)$, $i = 1, \ldots, n$, $S_k = X_1 + \ldots + X_k$, $k = 1, \ldots, n$. Показать, что $S_1, \ldots, S_n$ образуют марковскую цепь, выписать переходные плотности.

[По формуле преобразования плотности 3.26, учитывая, что якобиан перехода от X_i, $i=1, \ldots, n$, к S_k, $k=1, \ldots, n$, равен 1, находим

$$f_{S_1, \ldots, S_n}(s_1, \ldots, s_n) = f_{X_1, \ldots, X_n}(s_1, s_2 - s_1, \ldots, s_n - s_{n-1}) =$$
$$= f_{X_1}(s_1) f_{X_2}(s_2 - s_1) \ldots f_{X_n}(s_n - s_{n-1}).]$$

5.20. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и имеют стандартную экспоненциальную плотность. Используя 4.17 и 5.19, показать, что порядковые статистики $X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)}$ образуют цепь Маркова.

$$\left[X_{(k)} = \sum_{i=0}^{k} (X_{(i)} - X_{(i-1)}),\ X_0 \equiv 0, \text{ где слагаемые независимы.}\right]$$

§ 6.*

ПРОСТРАНСТВО И МЕРА

Характерная особенность современной математики — использование аксиоматического подхода для определения объекта, служащего предметом изучения. В теории вероятностей этот подход был разработан А. Н. Колмогоровым. Математическая структура, изучаемая в теории вероятностей, представляет собой тройку $(\Omega, \mathscr{A}, \mathbf{P})$, называемую *вероятностным пространством*. Здесь Ω — произвольное множество, точки которого называют *элементарными событиями*; $\mathscr{A}$ — класс подмножеств Ω, называемых *событиями*, удовлетворяющий требованиям:

1. $\varnothing, \Omega \in \mathscr{A}$;
2. $A \in \mathscr{A} \Rightarrow \bar{A} \in \mathscr{A}$;
3. $A_i \in \mathscr{A},\ i = 1, 2, \ldots, \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$;

$\mathbf{P}$ — функция на $\mathscr{A}$ со значениями $0 \leqslant \mathbf{P}(A) \leqslant 1$, $\mathbf{P}(\Omega) = 1$, $\mathbf{P}(\varnothing) = 0$, удовлетворяющая свойству *счетной аддитивности* (или σ-аддитивности)

$$A_i \in \mathscr{A},\ i = 1, 2, \ldots,\ A_i A_j = \varnothing,\ i \neq j \Rightarrow \mathbf{P}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathbf{P}(A_i)$$

и называемая *вероятностью (вероятностной мерой, распределением вероятностей)*.

Класс множеств $\mathscr{A}$, удовлетворяющий требованиям 1—3, называют *σ-алгеброй множеств*. Так как

$$\bigcap_{i=1}^{\infty} A_i = \overline{\bigcup_{i=1}^{\infty} \bar{A}_i},$$

то σ-алгебра замкнута относительно операции счетного пересечения. Полагая в свойстве счетной аддитивности $A_{n+1} = A_{n+2} = \ldots$

---

*) За исключением вводной части, этот параграф может быть опущен читателем, которого математические «тонкости» не интересуют.

$\ldots=\varnothing$, выводим из него свойство *конечной аддитивности* (или просто аддитивности)

$$A_i \in \mathcal{A},\ i=1,\ 2,\ \ldots,\ n,\ A_iA_j=\varnothing,\ i\neq j \Rightarrow \mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n}\mathbf{P}(A_i).$$

Всем перечисленным требованиям удовлетворяет изучавшееся в предыдущих разделах дискретное вероятностное пространство: Ω — не более чем счетное множество, $\mathcal{A}$ — класс всех подмножеств Ω,

$$\mathbf{P}(A)=\sum_{\omega\in A} p(\omega),\ A\in\mathcal{A},$$

где $0\leqslant p(\omega)\leqslant 1$ — произвольная функция на Ω, такая, что

$$\sum_{\omega\in\Omega} p(\omega)=1.$$

Вероятностные модели с несчетным множеством элементарных событий, которые рассматривались до сих пор, удовлетворяли более слабым требованиям: класс событий $\mathcal{A}$ представлял собой *алгебру множеств*, т. е. вместо условия 3 выполнялось

$$3a.\ A_i\in\mathcal{A},\ i=1,\ 2,\ \ldots,\ n,\Rightarrow \bigcup_{i=1}^{n} A_i\in\mathcal{A},$$

а мера $\mathbf{P}$ на алгебре множеств $\mathcal{A}$ удовлетворяла свойству конечной аддитивности. Недостатки такой модели отмечались неоднократно. В конкретных примерах возникала потребность в множествах элементарных событий вида $A=\bigcup_{i=1}^{\infty} A_i$, где $A_1\subseteq A_2\subseteq\ldots$ принадлежали алгебре $\mathcal{A}$, в то время как $A\overline{\in}\mathcal{A}$, или в множествах вида $B=\bigcup_{i=1}^{\infty} B_i\overline{\in}\mathcal{A}$, где $B_i\in\mathcal{A}$, $i=1,\ 2,\ \ldots,\ B_iB_j=\varnothing$ при $i\neq j$ (см., например, задачу о разорении игрока 2.19). Из затруднения мы выходили, *доопределяя* меру $\mathbf{P}$ по «непрерывности» или «счетной аддитивности»:

$$\mathbf{P}(A)=\lim_{n\to\infty}\mathbf{P}(A_n),\ \mathbf{P}(B)=\sum_{n=1}^{\infty}\mathbf{P}(B_n).$$

Оба этих соотношения — законные средства *вычисления* вероятностей событий при аксиоматическом подходе. При этом, однако, возникает другое затруднение — вопрос о существовании вероятностной модели, которая, с одной стороны, удовлетворяет аксиомам, а с другой — адекватно описывает тот реальный или умозрительный вероятностный эксперимент, для которого строится вероятностная модель. Поскольку построение вероятностной модели начинается с определения вероятностной меры $\mathbf{P}$ на некотором достаточно узком классе множеств элементарных событий, задача заключается в том, чтобы расширить область определения меры до некоторой σ-алгебры, на которой она была бы счетно-аддитивна. Проблема такого расширения, или, как говорят, *продолжения* меры, имеет общематематическое значение. Требование конечности меры всего пространства, как это имеет место для вероятностной меры, не является здесь принципиальным, и, если не оговорено противное, для мер будут допускаться и бесконечные значения.

Итак, рассмотрим функцию множества $\mu(A)$, $0 \leqslant \mu(A) \leqslant \infty$, заданную на некотором классе подмножеств $\mathscr{D}$ пространства Ω, от которого потребуем, как минимум, чтобы он являлся *полуалгеброй множеств*:

1. $\varnothing, \Omega \in \mathscr{D}$,
2. $A, B \in \mathscr{D} \Rightarrow AB \in \mathscr{D}$,
3. $A \in \mathscr{D} \Rightarrow \bar{A} = \bigcup\limits_{i=1}^{n} A_i,\ A_i \in \mathscr{D},\ A_i A_j = \varnothing,\ i \neq j$,

т. е. дополнение к любому элементу $A \in \mathscr{D}$ представимо в виде конечного объединения непересекающихся множеств из $\mathscr{D}$, а от функции $\mu(A)$ потребуем выполнения свойства аддитивности конечной или счетной: если $A_i \in \mathscr{D}$, $i=1, \ldots, n$, $n \leqslant \infty$, $\bigcup\limits_{i=1}^{n} A_i = A \in \mathscr{D}$, то $\mu(A) = \sum\limits_{i=1}^{n} \mu(A_i)$. Такие функции $\mu(A)$ будут называться конечно-аддитивными мерами, если в последнем соотношении допускаются конечные n, и счетно-аддитивными мерами, если допускается также $n = \infty$.

Хорошо известные примеры мер: длина, площадь, объем и, в случае подмножеств n-мерного пространства R^n, n-мерный объем. За первоначальную область определения этих мер можно принять совокупность всех промежутков $I_{a,b} = (a, b]$, $[a, b)$, (a, b), $[a, b]$, $-\infty \leqslant a \leqslant b \leqslant \infty$, прямоугольников $I_{a,b} \times I_{c,d}$, n-мерных параллелепипедов $I_{a_1,b_1} \times \ldots \times I_{a_n,b_n}$, соответственно полагая

$$\mu(I_{a_1,b_1} \times \ldots \times I_{a_n,b_n}) = (b_1 - a_1) \ldots (b_n - a_n),$$

где $-\infty \leqslant a_i \leqslant b_i \leqslant \infty$, $i=1, \ldots, n$. Легко видеть, что класс n-мерных параллелепипедов образует полуалгебру, а μ — аддитивная мера на этой полуалгебре (см. 6.4).

6.1. Задача. Пусть $\mathscr{D}$ есть полуалгебра множеств. Рассмотрим класс множеств $\mathscr{A}$, состоящий из всевозможных конечных объединений непересекающихся элементов из $\mathscr{D}$. Показать, что $\mathscr{A}$ является алгеброй множеств.

[Пусть $A=\bigcup_{\{i\}}A_i$, $B=\bigcup_{\{j\}}B_j$ — конечные объединения непересекающихся элементов из $\mathscr{D}$. Тогда

$$AB=\bigcup_{\{i,j\}}A_iB_j$$

— конечное объединение непересекающихся элементов A_iB_j из $\mathscr{D}$ и, следовательно, $AB\in\mathscr{A}$. Далее,

$$\bar{A}=\overline{\bigcup_{\{i\}}A_i}=\bigcap_{\{i\}}\bar{A}_i.$$

Поскольку $\bar{A}_i$ по определению полуалгебры $\mathscr{D}$ представимо в виде конечного объединения непересекающихся элементов из $\mathscr{D}$, то $\bar{A}_i\in\mathscr{A}$, $i=1, \ldots, n$. По уже доказанному $\mathscr{A}$ замкнуто относительно образования пересечений и, следовательно, $\bar{A}\in\mathscr{A}$. Остается заметить, что $A\bigcup B=\overline{\bar{A}\bar{B}}\in\mathscr{A}$.]

Алгебра $\mathscr{A}$, построенная по полуалгебре $\mathscr{D}$ в 6.1, содержит, очевидно, любые конечные объединения элементов из $\mathscr{D}$ и является минимальной алгеброй множеств, содержащей $\mathscr{D}$. Ее называют алгеброй, *порожденной* полуалгеброй $\mathscr{D}$.

6.2. Задача. Пусть $\mathscr{D}$ — полуалгебра множеств, $\mathscr{A}$ — порожденная ею алгебра, μ — мера на $\mathscr{D}$. Для любого $A\in\mathscr{A}$,

$$A=\bigcup_{i=1}^{n}A_i,\ A_i\in\mathscr{D},\ i=1, \ldots, n,\ A_iA_j=\varnothing,\ i\neq j,$$ положим

$$\mu(A)=\sum_{i=1}^{n}\mu(A_i).$$

Показать, что этой формулой корректно определена функция $\mu(A)$ на $\mathscr{A}$, которая является конечно-аддитивной мерой. Если мера μ на $\mathscr{D}$ счетно-аддитивна, то это же верно для меры μ на $\mathscr{A}$. [Допустим, что $A=\bigcup_{\{i\}}A_i$ и $A=\bigcup_{\{j\}}B_j$ — два различных представления множества $A\in\mathscr{A}$ в виде конечного объединения непересекающихся элементов из $\mathscr{D}$. В таком случае при любом i имеем

$$A_i=A_iA=\bigcup_{\{j\}}A_iB_j,$$

а из аддитивности меры μ на $\mathscr{D}$ получаем

$$\mu(A_i)=\Sigma_{\{j\}}\mu(A_iB_j).$$

Аналогично при любом j

$$\mu(B_j)=\Sigma_{\{i\}}\mu(B_jA_i).$$

Таким образом,

$$\Sigma_{\{i\}}\mu(A_i)=\Sigma_{\{i,j\}}\mu(A_iB_j)=\Sigma_{\{j\}}\mu(B_j).$$

Докажем аддитивность (σ-аддитивность) функции μ на $\mathscr{A}$. Пусть $A=\bigcup_{\{i\}}A_i$, где A, $A_i\in\mathscr{A}$, $A_iA_j=\varnothing$ при $i\neq j$, а индекс i пробегает некоторое конечное (счетное) множество. Представим каждое A_i в виде конечного объединения непересекающихся элементов из $\mathscr{D}$: $A_i=\bigcup_{(j)}A_{ij}$, $A_{ij}\in\mathscr{D}$, и воспользуемся аддитивностью (σ-аддитивностью) меры μ на $\mathscr{D}$:

$$\mu(A)=\mu(\bigcup_{(i,j)}A_{ij})=\Sigma_{(i,j)}\mu(A_{ij})=\Sigma_{(i)}\mu(A_i).]$$

6.3.* Задача. Рассмотрим полуалгебру всех промежутков $I_{a,b}=(a, b)$, $(a, b]$, $[a, b)$, $[a, b]$, $-\infty\leqslant a\leqslant b\leqslant\infty$. Для любого конечного промежутка $I_{a,b}$ положим $\mu(I_{a,b})=b-a$, а для промежутка, у которого хотя бы один конец бесконечен, положим $\mu(I)=\infty$. Показать, что μ — счетно-аддитивная мера на $\mathscr{D}$, а ввиду 6.2 и на порожденной $\mathscr{D}$ алгебре $\mathscr{A}$.

[Пусть $-\infty<a<b<\infty$, $I_{a,b}=\bigcup_{k=1}^{n} I_{a_k,b_k}$, где промежутки I_{a_k,b_k}, $k=1, 2, \ldots$, попарно не пересекаются. Очевидно, при любом конечном n

$$I_{a,b}\supseteq\bigcup_{k=1}^{n} I_{a_k,b_k},\quad b-a\geqslant\sum_{k=1}^{n}(b_k-a_k).$$

Устремляя n к ∞, получаем

$$b-a\geqslant\sum_{k=1}^{\infty}(b_k-a_k).$$

Для вывода неравенства в противоположную сторону возьмем произвольное $\varepsilon>0$ и образуем интервалы $(a_k', b_k')=(a_k-\varepsilon 2^{-k}, b_k+\varepsilon 2^{-k})$, $k=1, 2, \ldots$. Очевидно,

$$[a, b]\subseteq\bigcup_{k=1}^{\infty}(a_k', b_k').$$

По лемме Гейне — Бореля выделим конечное покрытие отрезка $[a, b]$:

$$[a, b]\subseteq\bigcup_{k=1}^{m}(a_k', b_k'),$$

откуда

$$b-a\leqslant\sum_{k=1}^{m}(b_k'-a_k')\leqslant\sum_{k=1}^{\infty}(b_k'-a_k')=\sum_{k=1}^{\infty}(b_k-a_k)+\varepsilon.$$

Ввиду произвольности ε получаем

$$b-a\leqslant\sum_{k=1}^{\infty}(b_k-a_k),$$

что завершает проверку счетной аддитивности для конечных промежутков.

Разберем случай бесконечного промежутка

$$I_{a,b}=\bigcup_{k=1}^{\infty} I_{a_k,b_k}.$$

Для любого отрезка $[a', b']\subseteq I_{a,b}$ имеем

$$[a', b']\subseteq \bigcup_{k=1}^{\infty} I_{a_k,b_k},$$

откуда тем же методом, что и выше, выводится неравенство

$$b'-a'\leqslant \sum_{k=1}^{\infty}(b_k-a_k).$$

Остается заметить, что отрезок $[a', b']$ может быть взят произвольно большой длины.]

6.4.* Задача. Обобщить 6.3 на n-мерные параллелепипеды. [Решение проводится по схеме 6.3. Дальнейшее обобщение см. в 6.24—6.26.]

6.5. Задача. Показать, что любая мера на алгебре множеств $\mathcal{A}$ обладает следующими свойствами:

(I) если $A\subseteq B$, A, $B\in\mathcal{A}$ и $\mu(B)<\infty$, то $\mu(B\bar{A})=\mu(B)-\mu(A)$;

(II) если $A\subseteq B$, A, $B\in\mathcal{A}$, то $\mu(A)\leqslant\mu(B)$;

(III) если $A_i\in\mathcal{A}$, $i=1, \ldots, n$, то $\mu\left(\bigcup_{i=1}^{n} A_i\right)\leqslant \sum_{i=1}^{n}\mu(A_i)$;

(IV) если мера μ счетно-аддитивна на $\mathcal{A}$, то (III) выполняется с заменой n на ∞.

Свойство (II) называют *монотонностью* меры μ, а свойства (III), (IV) — *полуаддитивностью*.

[(I) Так как $\mu(A)+\mu(\bar{A}B)=\mu(B)$, то $\mu(A)$, $\mu(\bar{A}B)<\infty$ и $\mu(\bar{A}B)=\mu(B)-\mu(A)$. Свойство (II) тривиально при $\mu(B)=\infty$ и вытекает из (I) при $\mu(B)<\infty$. (III) Положим $B_1=A_1$, $B_i=A_i\bar{A}_{i-1}\ldots\bar{A}_1$, $i=2, \ldots, n$, так что $B_iB_j=\varnothing$ при $i\neq j$ и $\bigcup_{i=1}^{n} B_i=\bigcup_{i=1}^{n} A_i$. Поэтому

$$\mu\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n}\mu(B_i)\leqslant\sum_{i=1}^{n}\mu(A_i),$$

где на последнем шаге использовано свойство (II) для $B_i\subseteq A_i$. Свойство (IV) устанавливается аналогично.]

Последовательность множеств A_i, $i=1, 2, \ldots$, называется *возрастающей (убывающей)*, если $A_i\subseteq A_{i+1}$ $(A_i\supseteq A_{i+1})$, $i=1, 2, \ldots$. Возрастающие и убывающие последовательности множеств называются *монотонными*. Мера μ на алгебре множеств $\mathcal{A}$ называется *непрерывной снизу (сверху)*, если для любой возрастающей (убывающей) последовательности множеств A_i, $i=1, 2, \ldots$, такой, что

$$A=\bigcup_{i=1}^{\infty} A_i \in \mathscr{A} \, (A=\bigcap_{i=1}^{\infty} A_i \in \mathscr{A},\ \mu(A_1)<\infty),$$

$$\mu(A)=\lim_{n\to\infty}\mu(A).$$

Если мера μ непрерывна сверху и снизу, то она называется *непрерывной*. Для конечных мер из соотношения двойственности $\overline{\bigcup A_i}=\bigcap \overline{A_i}$ вытекает, что мера непрерывна, если она непрерывна сверху (снизу). Мера μ называется непрерывной в $\varnothing$, если для любой убывающей к $\varnothing$ последовательности $A_i \in \mathscr{A}$, $\mu(A_i)<\infty$, $i=1, 2, \ldots$, имеет место соотношение

$$\lim_{n\to\infty}\mu(A_n)=0.$$

6.6. Задача. Если мера μ на алгебре множеств $\mathscr{A}$ непрерывна в $\varnothing$, то она непрерывна.

[Пусть $A_i \in \mathscr{A}$, $i=1, 2, \ldots$, — убывающая последовательность, $\mu(A_1)<\infty$, $A=\bigcap_{i=1}^{\infty} A_i \in \mathscr{A}$. Тогда $B_i=A_i\bar{A}$, $i=1, 2, \ldots$, образуют убывающую к пустому множеству последовательность и

$$\mu(B_n)=\mu(A_n)-\mu(A)\to 0, \quad n\to\infty.$$

Аналогично устанавливается непрерывность меры μ снизу.]

6.7. Задача. Показать, что: (I) счетно-аддитивная мера μ на алгебре множеств $\mathscr{A}$ является непрерывной; (II) непрерывная снизу конечно-аддитивная мера на $\mathscr{A}$ счетно-аддитивна; (III) непрерывная в $\varnothing$, конечная конечно-аддитивная мера на $\mathscr{A}$ счетно-аддитивна.

[(I) Пусть $A_1 \supseteq A_2 \supseteq \ldots$, $A_i \in \mathscr{A}$, $i=1, 2, \ldots$ $\mu(A_1)<\infty$ $\bigcap_{i=1}^{\infty} A_i=\varnothing$. Положим $B_i=A_i\bar{A}_{i+1}$, $i=1, 2, \ldots$, так что при $i<j$, $B_iB_j=A_iA_j\overline{A_{i+1}}\,\overline{A_{j+1}}=A_j\overline{A_{i+1}}\subseteq A_j\bar{A}_j=\varnothing$, $\bigcup_{i=1}^{\infty} B_i=A_1$. Предполагая, что мера μ счетно-аддитивна, получаем

$$\mu(A_1)=\lim_{n\to\infty}\sum_{i=1}^{n}\mu(B_i)=\lim_{n\to\infty}\sum_{i=1}^{n}(\mu(A_i)-\mu(A_{i+1}))=$$
$$=\lim_{n\to\infty}(\mu(A_1)-\mu(A_{n+1}))=\mu(A_1)-\lim_{n\to\infty}\mu(A_n),$$

т. е. $\lim_{n\to\infty}\mu(A_n)=0$ и мера μ непрерывна.

(II) Для любых $A_i \in \mathscr{A}$, $i=1, 2, \ldots$, $A_iA_j=\varnothing$ при $i\neq j$, $\bigcup_{i=1}^{\infty} A_i=$

$=A\in\mathscr{A}$ образуем возрастающую последовательность множеств $B_n=\bigcup_{i=1}^{n} A_i$, $n=1, 2, \ldots$, $\bigcup_{n=1}^{\infty} B_n=\bigcup_{i=1}^{\infty} A_i$. По непрерывности снизу

$$\mu(A)=\mu\Big(\bigcup_{n=1}^{\infty} B_n\Big)=\lim_{n\to\infty}\mu(B_n)=\lim_{n\to\infty}\sum_{i=1}^{n}\mu(A_i),$$

если $\mu(A_i)<\infty$, $i=1, 2, \ldots$. Если при некотором n $\mu(A_n)=\infty$, то из $A_n\subseteq A$ вытекает, что $\mu(A)=\infty$, что и завершает доказательство счетной аддитивности.

(III) В обозначениях (II) имеем

$$\mu(A)=\sum_{i=1}^{n}\mu(A_i)+\mu\Big(\bigcup_{i=n+1}^{\infty} A_i\Big).$$

Поскольку $\bigcap_{n=1}^{\infty}\bigcup_{i=n+1}^{\infty} A_i=\varnothing$, то второе слагаемое в последнем равенстве стремится к 0 при $n\to\infty$, что и требовалось установить.]

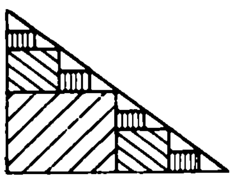

Рис. 22

Перейдем к вопросу о продолжении счетно-аддитивной меры с алгебры множеств $\mathscr{A}$ на более обширный класс множеств. Рассматривая для примера алгебру конечных объединений непересекающихся параллелепипедов с n-мерным объемом в качестве меры (см. 6.4), заметим, что класс «измеримых» подмножеств можно расширить, образуя счетные объединения непересекающихся параллелепипедов. Так, на рис. 22 показано, как может быть «получена» при этом площадь прямоугольного треугольника: $\mu(B)=\sum_{i=1}^{\infty}\mu(B_i)$.

Поскольку конечное объединение прямоугольников $\bigcup_{i=1}^{n} B_i=A_n$ есть элемент порожденной полуалгеброй алгебры, то указанный процесс расширения можно иначе определить так: для любой возрастающей последовательности $A_i\in\mathscr{A}$, $i=1, 2, \ldots$, положим

$$\mu(A)=\lim_{n\to\infty}\mu(A_n),\quad A=\bigcup_{n=1}^{\infty} A_n.$$

6.8.* Задача. Показать, что для любых двух возрастающих последовательностей A_i, $i=1, 2, \ldots$, A_j', $j=1, 2, \ldots$, множеств из алгебры $\mathscr{A}$, таких, что $\bigcup_{i=1}^{\infty} A_i\subseteq\bigcup_{j=1}^{\infty} A_j'$, и счетно-аддитивной меры μ имеет место неравенство

$$\lim_{n\to\infty}\mu(A_n)\leqslant\lim_{n\to\infty}\mu(A_n').$$

Это неравенство переходит в равенство, если $\bigcup\limits_{i=1}^{\infty} A_i = \bigcup\limits_{j=1}^{\infty} A_j'$. [Поскольку для любого i последовательность A_iA_j', $j=1, 2, \ldots$, возрастает и $\bigcup\limits_{j=1}^{\infty} A_iA_j' = A_i(\bigcup\limits_{j=1}^{\infty} A_j') = A_i$, то, воспользовавшись непрерывностью меры μ на $\mathscr{A}$, получаем в предположении, что $\mu(A_iA_j') < \infty$:

$$\lim_{n\to\infty} \mu(A_mA_n') = \mu(A_m).$$

Так как $A_mA_n' \subseteq A_n'$, то $\mu(A_mA_n') \leqslant \mu(A_n')$ и

$$\lim_{n\to\infty} \mu(A_n') \geqslant \mu(A_m).$$

Остается перейти в этом неравенстве к пределу при $m\to\infty$. В случае, когда при некоторых i, j $\mu(A_iA_j') = \infty$, имеем $\mu(A_i) = \infty$, $\mu(A_j') = \infty$, и утверждение задачи тривиально.]

Из 6.8 вытекает, что счетно-аддитивная мера μ может быть продолжена с алгебры $\mathscr{A}$ на класс множеств $\mathscr{U} = \mathscr{U}(\mathscr{A})$, образованный из всевозможных счетных объединений $\bigcup\limits_{i=1}^{\infty} A_i$ элементов из $\mathscr{A}$ (или, что эквивалентно, из счетных объединений возрастающих последовательностей из $\mathscr{A}$), если положить

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mu(\bigcup_{i=1}^{n} A_i).$$

6.9.* Задача. Показать, что выше определенный класс множеств $\mathscr{U} = \mathscr{U}(\mathscr{A})$ замкнут относительно операций конечного пересечения и счетного объединения.

[Пусть $A_{i,1}$, $i=1, 2, \ldots$, $A_{j,2}$, $j=1, 2, \ldots$, — две возрастающие последовательности элементов из алгебры $\mathscr{A}$, порождающей класс $\mathscr{U}$, $A_1 = \bigcup\limits_{i=1}^{\infty} A_{i,1}$, $A_2 = \bigcup\limits_{j=1}^{\infty} A_{j,2}$. Полагая для краткости $i \vee j = \max(i, j)$, имеем

$$A_1A_2 = \bigcup_{(i,j)} A_{i,1}A_{j,2} \subseteq \bigcup_{(i,j)} A_{i\vee j,1}A_{i\vee j,2} = \bigcup_{(n)} A_{n,1}A_{n,2}.$$

Очевидно, что

$$A_1A_2 \supseteq \bigcup_{n=1}^{\infty} A_{n,1}A_{n,2}$$

и потому

$$A_1A_2 = \bigcup_{n=1}^{\infty} A_{n,1}A_{n,2}.$$

Поскольку последовательность $A_{n,1}A_{n,2}$, $n=1, 2, \ldots$, возрастает и $A_{n,1}A_{n,2} \in \mathscr{A}$, то $A_1A_2 \in \mathscr{U}$. Точно так же устанавливается, что $A_1 \cup A_2 \in \mathscr{A}$.

Покажем, что если A_n, $n=1, 2, \ldots$, — возрастающая последовательность элементов из $\mathscr{U}$, то $A=\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{U}$. Пусть $A_n = \bigcup\limits_{m=1}^{\infty} A_{n,m}$, где $A_{n,m}$, $m=1, 2, \ldots$, — возрастающая последовательность элементов из $\mathscr{A}$ при каждом фиксированном n. Имеем

$$A=\bigcup_{n=1}^{\infty}\bigcup_{m=1}^{\infty} A_{n,m}=\bigcup_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_{n,m}=\bigcup_{m=1}^{\infty}\bigcup_{n=1}^{m} A_{n,m},$$

причем

$$\bigcup_{n=1}^{m} A_{n,m} \subseteq \bigcup_{n=1}^{m+1} A_{n,m} \subseteq \bigcup_{n=1}^{m+1} A_{n,m+1},$$

где мы неоднократно использовали возрастание последовательности $A_{n,m}$, $m=1, 2, \ldots$. Таким образом, A представлено в виде объединения возрастающей последовательности элементов из $\mathscr{A}$ и, следовательно, входит в $\mathscr{U}$. Для произвольной последовательности $A_n \in \mathscr{U}$, $n=1, 2, \ldots$, перейдем к возрастающей последовательности $B_n = \bigcup\limits_{i=1}^{n} A_i$, и, учитывая, что $\mathscr{U}$ замкнуто относительно операции конечного объединения, заключаем, что $\bigcup\limits_{i=1}^{\infty} A_i = \bigcup\limits_{n=1}^{\infty} B_n \in \mathscr{U}$].

6.10.* Задача. Показать, что функция μ на классе множеств $\mathscr{U}(\mathscr{A})$ аддитивна, непрерывна снизу и, следовательно, счетно-аддитивна. Кроме того, $\mu(A) \leqslant \mu(B)$, если $A \subseteq B$, A, $B \in \mathscr{U}$.

[Последнее свойство доказано в 6.8. Далее, пусть A_1, $A_2 \in \mathscr{U}$, $A_1 A_2 = \varnothing$. В обозначениях решения задачи 6.9 имеем

$$A_1 \bigcup A_2 = \bigcup_{n=1}^{\infty} (A_{n,1} \bigcup A_{n,2}), \quad A_{n,1} A_{n,2} = \varnothing,$$

откуда

$$\mu(A_1 \bigcup A_2) = \lim_{n\to\infty} \mu(A_{n,1} \bigcup A_{n,2}) = \lim_{n\to\infty} (\mu(A_{n,1}) + \mu(A_{n,2})) =$$
$$= \mu(A_1) + \mu(A_2).$$

Снова используя 6.9, возьмем возрастающую последовательность $A_n \in \mathscr{U}$, $n=1, 2, \ldots$, $A=\bigcup\limits_{n=1}^{\infty} A_n$, $A_n = \bigcup\limits_{m=1}^{\infty} A_{n,m}$, где $A_{n,m}$, $m=1, 2, \ldots$, — возрастающая последовательность элементов из $\mathscr{A}$, и запишем равенство

$$A = \bigcup_{m=1}^{\infty} A_m = \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{m} A_{n,m},$$

где $\bigcup\limits_{n=1}^{m} A_{n,m}$, $m=1, 2, \ldots$,— возрастающая последовательность элементов из $\mathscr{A}$. Поэтому

$$\mu(A)=\lim_{m\to\infty}\mu\Big(\bigcup_{n=1}^{m} A_{n,m}\Big).$$

С другой стороны,

$$\bigcup_{n=1}^{m} A_{n,m}\subseteq\bigcup_{n=1}^{m} A_n=A_m,$$

откуда

$$\mu\Big(\bigcup_{n=1}^{m} A_{n,m}\Big)\leqslant\mu(A_m),$$

и, переходя к пределу при $m\to\infty$ и учитывая монотонность последовательности $\mu(A_m)$, получаем

$$\mu(A)\leqslant\lim_{m\to\infty}\mu(A_m).$$

Противоположное неравенство следует из того, что $A_m\subseteq A$.]

Класс множеств $\mathscr{U}$ с определенной на нем мерой μ удовлетворяет многим практическим потребностям. Вместе с тем операция счетного пересечения выводит за пределы $\mathscr{U}$. К тому же при $A\subseteq B$, A, $B\in\mathscr{U}$, вообще говоря, $B\bar{A}\overline{\in}\mathscr{U}$. Можно было бы доопределить μ, полагая $\mu(B\bar{A})=\mu(B)-\mu(A)$ при $A\subseteq B$, A, $B\in\mathscr{U}$, $\mu(B)<\infty$, и

$$\mu\Big(\bigcap_{i=1}^{\infty} A_i\Big)=\lim_{n\to\infty}\mu\Big(\bigcap_{i=1}^{n} A_i\Big),\quad A_i\in\mathscr{U},\quad i=1, 2, \ldots,$$

но, к сожалению, этим все не кончается, поскольку для расширенного класса операция счетного объединения, вообще говоря, выводит за его пределы. Вместо того чтобы продолжать этот бесконечный процесс, используем множества класса $\mathscr{U}$ для аппроксимации подмножеств Ω общего вида. Для любого $A\subseteq\Omega$ определим *внешнюю меру*

$$\mu^*(A)=\inf\{\mu(B): A\subseteq B,\ B\in\mathscr{U}\}$$

как точную нижнюю грань мер $\mu(B)$ всевозможных элементов B из $\mathscr{U}$, содержащих A. Очевидно, $\mu^*(A)=\mu(A)$ для $A\in\mathscr{U}$. Из монотонности меры μ на $\mathscr{U}$ вытекает, что $\mu^*(A_1)\leqslant\mu^*(A_2)$, если $A_1\subseteq A_2$.

6.11. Задача. Показать, что для конечной меры μ

$$\mu^*\big(A_1\bigcup A_2\big)\leqslant\mu_1^*(A_1)+\mu_2^*(A_2)-\mu^*(A_1A_2).$$

[Для любого $\varepsilon>0$ подберем такие $B_1\supseteq A_1$, $B_2\supseteq A_2$, B_1, $B_2\in\mathscr{U}$, что

$$\mu^*(A_i)\geqslant\mu(B_i)-\varepsilon,\quad i=1, 2.$$

Учитывая, что $A_1 \bigcup A_2 \subseteq B_1 \bigcup B_2$, $A_1A_2 \subseteq B_1B_2$, имеем

$$\mu\ (A_1 \bigcup A_2) \leqslant \mu\ (B_1 \bigcup B_2) = \mu(B_1) + \mu(B_2) - \mu(B_1B_2) \leqslant$$
$$\leqslant \mu^*(A_1) + \varepsilon + \mu^*(A_2) + \varepsilon - \mu^*(A_1A_2).$$

Ввиду произвольности ε получаем требуемое.]

6.12.* Задача. Показать, что при конечной мере μ внешняя мера непрерывна снизу.

[Пусть $A_n \subseteq \Omega$, $n=1, 2, \ldots$, — возрастающая последовательность, $A = \bigcup\limits_{n=1}^{\infty} A_n$. Для любого $\varepsilon>0$ выберем $B_n \in \mathcal{U}$, $B_n \supseteq A_n$, такое, что

$$\mu(B_n) \leqslant \mu^*(A_n) + \varepsilon 2^{-n}, \quad n=1, 2, \ldots .$$

Докажем индукцией по n, что

$$\mu\ (\bigcup_{i=1}^{n} B_i) \leqslant \mu^*(A_n) + \varepsilon(1-2^{-n}).$$

Имеем, учитывая свойство монотонности внешней меры μ^*,

$$\mu\ (\bigcup_{i=1}^{n+1} B_n) = \mu\ (\bigcup_{i=1}^{n} B_i) + \mu(B_{n+1}) - \mu\ (\bigcup_{i=1}^{n} B_iB_{n+1}) \leqslant$$

$$\leqslant \mu^*(A_n) + \varepsilon(1-2^{-n}) + \mu^*(A_{n+1}) + \varepsilon 2^{-n-1} - \mu^*(\bigcup_{i=1}^{n} A_iA_{n+1}) =$$

$$= \mu^*(A_n) + \varepsilon(1-2^{-n-1}) + \mu^*(A_{n+1}) - \mu^*(A_{n+1}) = \mu^*(A_n) + \varepsilon(1-2^{-n-1}).$$

Воспользовавшись доказанным неравенством, запишем

$$\mu\ (\bigcup_{i=1}^{n} B_i) - \varepsilon(1-2^{-n}) \leqslant \mu^*(A_n) \leqslant \mu^*(A) \leqslant \mu\ (\bigcup_{i=1}^{\infty} B_i).$$

Переходя к пределу по n, получаем

$$\mu\ (\bigcup_{i=1}^{\infty} B_i) - \varepsilon \leqslant \lim_{n\to\infty} \mu^*(A_n) \leqslant \mu^*(A) \leqslant \mu\ (\bigcup_{i=1}^{\infty} B_i).$$

Ввиду произвольности ε заключаем, что

$$\lim_{n\to\infty} \mu^*(A_n) = \mu^*(A).]$$

На множестве всех подмножеств Ω внешняя мера μ^* полуаддитивна (см. 6.11). Естественно попытаться сузить класс подмножеств, чтобы сохранить свойство аддитивности меры. Рассмотрим вначале случай конечной меры μ и, не ограничивая общности, предположим, что $\mu(\Omega)=1$. Частным проявлением аддитивности внешней меры μ^* в этом случае является равенство

$$\mu^*(A) + \mu^*(\bar{A}) = 1.$$

Всякое подмножество $A \subseteq \Omega$, удовлетворяющее указанному условию, назовем μ^*-*измеримым* и обозначим $\mathscr{B}^*$ класс всех μ^*-измеримых подмножеств Ω. Отметим, что для произвольного подмножества $A \subseteq \Omega$ имеем (см. 6.11)

$$1 = \mu(\Omega) = \mu^*(\Omega) = \mu^*(A \cup \bar{A}) \leqslant \mu^*(A) + \mu^*(\bar{A}).$$

Из последующего будет ясно, почему μ^* следует считать продолжением меры μ с алгебры множеств $\mathscr{A}$ на $\mathscr{B}^*$. Однако уже сейчас стоит обратить внимание на естественность *определения* меры μ на множествах из $\mathscr{B}^*$ как внешней меры этих множеств. Напомним, что класс множеств $\mathscr{U}$, служащий для определения внешней меры, не замкнут, вообще говоря, относительно перехода к дополнению. Его можно было бы расширить, положив $\mu(\bar{A}) = 1 - \mu(A)$, $A \in \mathscr{U}$. Допустив это расширение, возьмем произвольное множество $B \in \mathscr{B}^*$ и подберем по нему такие элементы

$A_1,\ A_2 \in \mathscr{U},\ A_1 \supseteq B,\ A_2 \supseteq \bar{B}$, что

$$\mu(A_1) \leqslant \mu^*(B) + \varepsilon/2, \quad \mu(A_2) \leqslant \mu^*(\bar{B}) + \varepsilon/2.$$

Тогда

$$1 = \mu(\Omega) = \mu(A_1 \cup A_2) \leqslant \mu(A_1) + \mu(A_2) \leqslant$$
$$\leqslant \mu^*(B) + \varepsilon/2 + \mu^*(\bar{B}) + \varepsilon/2 = 1 + \varepsilon,$$

или

$$1 - \mu(A_2) \leqslant \mu(A_1) \leqslant 1 - \mu(A_2) + \varepsilon.$$

Таким образом, для любого множества $B \in \mathscr{B}^*$ и любого $\varepsilon > 0$ найдутся такие множества $\bar{A}_2 \subseteq B \subseteq A_1$, A_1, $A_2 \in \mathscr{U}$, что

$$\mu(\bar{A}_2) \leqslant \mu(A_1) \leqslant \mu(\bar{A}_2) + \varepsilon,$$

т. е. разность между μ-мерами «описанного» вокруг B и «вписанного» в B множеств не превосходит ε.

6.13. Задача. Показать, что $\mathscr{U} \subseteq \mathscr{B}^*$.

[Пусть $A \in \mathscr{U}$, $A = \bigcup_{i=1}^{\infty} A_i$, $A_i \in \mathscr{A}$, $i = 1, 2, \ldots$, $A_i A_j = \varnothing$ при $i \neq j$, так что

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i).$$

Множество $\bigcup_{i=1}^{n} A_i \subseteq A$ принадлежит алгебре множеств $\mathscr{A}$,

$$\mu^*(\bar{A}) \leqslant \mu\left(\overline{\bigcup_{i=1}^{n} A_i}\right) = 1 - \mu\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \mu(A) + \sum_{i=n+1}^{\infty} \mu(A_i).$$

Выбирая n достаточно большим, можно сделать остаток ряда меньше любого наперед заданного $\varepsilon>0$. В таком случае

$$1 \leqslant \mu^*(A)+\mu^*(\bar{A}) \leqslant \mu(A)+1-\mu(A)+\varepsilon=1+\varepsilon,$$

откуда ввиду произвольности ε вытекает $\mu^*(A)+\mu^*(\bar{A})=1$, т. е. $A \in \mathscr{B}^*$.]

6.14*. Задача. Показать, что $\mathscr{B}^*$ есть алгебра множеств, а μ^* — конечно-аддитивная мера на $\mathscr{B}^*$.

[Из определения $\mathscr{B}^*$ вытекает, что вместе с $A \in \mathscr{B}^*$ также $\bar{A} \in \mathscr{B}^*$. Пусть $A, B \in \mathscr{B}^*$. Покажем одновременно, что $A \cup B$, $AB \in \mathscr{B}^*$. По формуле 6.11 имеем

$$\mu^*(A \cup B)+\mu^*(AB) \leqslant \mu^*(A)+\mu^*(B),$$

$$\mu^*(\bar{A} \cup \bar{B})+\mu^*(\bar{A}\bar{B}) \leqslant \mu^*(\bar{A})+\mu^*(\bar{B}).$$

Принимая во внимание, что $\overline{A \cup B}=\bar{A}\bar{B}$, $\overline{AB}=\bar{A} \cup \bar{B}$, запишем

$$1 \leqslant \mu^*(A \cup B)+\mu^*(\bar{A}\bar{B}), \quad 1 \leqslant \mu^*(AB)+\mu^*(\bar{A} \cup \bar{B}).$$

Поскольку $\mu^*(A)+\mu^*(\bar{A})=1$, $\mu^*(B)+\mu^*(\bar{B})=1$, то, складывая обе пары написанных выше неравенств, получаем

$$2 \leqslant \mu^*(A \cup B)+\mu^*(AB)+\mu^*(\bar{A} \cup \bar{B})+\mu^*(\bar{A}\bar{B}) \leqslant 2.$$

Это означает, что все неравенства обращаются в равенства. Таким образом,

$$\mu^*(A \cup B)+\mu^*(\overline{A \cup B})=1, \quad \mu^*(AB)+\mu^*(\overline{AB})=1,$$

т. е. $A \cup B$, $AB \in \mathscr{B}^*$, и

$$\mu^*(A \cup B)+\mu^*(AB)=\mu^*(A)+\mu^*(B),$$

т. е. μ^* аддитивна на $\mathscr{B}^*$.]

6.15. Задача. Показать, что алгебра множеств $\mathscr{B}^*$ замкнута относительно операции счетного объединения возрастающих последовательностей множеств, а следовательно, $\mathscr{B}^*$ замкнуто относительно счетного объединения любых ее элементов, т. е. является σ-алгеброй.

[Пусть $B_n \in \mathscr{B}^*$, $n=1, 2, \ldots$, — возрастающая последовательность, $B=\bigcup_{n=1}^{\infty} B_n$. Ввиду 6.12 $\mu^*(B)=\lim_{n\to\infty} \mu^*(B_n)$, а из монотонности μ^* вытекает, что

$$\mu^*(\bar{B})=\mu^*\Big(\bigcap_{n=1}^{\infty} \bar{B}_n\Big) \leqslant \mu^*(\bar{B}_m)$$

при любом m. Отсюда при $n\to\infty$ имеем

$$1=\lim_{n\to\infty}(\mu^*(B_n)+\mu^*(\bar{B}_n)) \geqslant \mu^*(B)+\mu^*(\bar{B}).$$

С другой стороны, всегда $\mu^*(B)+\mu^*(\bar{B})\geqslant 1$, что и доказывает утверждение.]

Соберем вместе утверждения, установленные в 6.12—6.15 для случая вероятностной меры $\mu(\mu(\Omega)=1)$. Показано, что если μ — вероятностная счетно-аддитивная мера на алгебре множеств $\mathscr{A}$, то существует непрерывная снизу конечно-аддитивная, а следовательно, и счетно-аддитивная (см. 6.7) мера μ^* на некоторой σ-алгебре $\mathscr{B}^*$, включающей $\mathscr{A}$, такая, что $\mu^*(A)=\mu(A)$ для всех $A\in\mathscr{A}$. Дополним это утверждение результатом задач 6.1—6.2 о том, что счетно-аддитивная мера на алгебре множеств $\mathscr{A}$ может быть получена продолжением счетно-аддитивной меры, заданной на полуалгебре множеств $\mathscr{D}$, порождающей алгебру $\mathscr{A}$.

6.16. Задача. Пусть $\mathscr{B}_\alpha$ — некоторое семейство σ-алгебр подмножеств из Ω. Показать, что $\bigcap_{(\alpha)}\mathscr{B}_\alpha$ также будет σ-алгеброй.

[Поскольку $\varnothing, \Omega\in\mathscr{B}_\alpha$ при всех α, то $\varnothing$, $\Omega\in\bigcap_{(\alpha)}\mathscr{B}_\alpha$. Если $B\in\bigcap_{(\alpha)}\mathscr{B}_\alpha$, т. е. $B\in\mathscr{B}_\alpha$ при всех α, то $\bar{B}\in\mathscr{B}_\alpha$ при всех α, т. е. $\bar{B}\in\bigcap_{(\alpha)}\mathscr{B}_\alpha$, и точно так же устанавливается, что класс множеств $\bigcap_{(\alpha)}\mathscr{B}_\alpha$ замкнут относительно операции счетного объединения.]

Рассмотрим семейство всех σ-алгебр $\mathscr{B}_\alpha$, которые содержат заданный класс подмножеств S. Это семейство непусто, так как система всех подмножеств из Ω образует σ-алгебру, включающую S. Следовательно, σ-алгебра $\bigcap_{(\alpha)}\mathscr{B}_\alpha$ — наименьшая σ-алгебра, содержащая S. Она называется также *σ-алгеброй, порожденной* S, и обозначается $\sigma(S)$.

6.17. Задача. Пусть μ — счетно-аддитивная вероятностная мера на алгебре множеств $\mathscr{A}$, μ^* — внешняя мера, построенная по μ, ν — некоторая счетно-аддитивная мера на $\sigma(\mathscr{A})$, совпадающая с μ на $\mathscr{A}$. Показать, что ν совпадает с μ^* на $\sigma(\mathscr{A})$, проверив последовательно, что: (I) на классе $\mathscr{U}(\mathscr{A})$ меры ν и μ^* совпадают; (II) $\nu\leqslant\mu^*$ на $\sigma(\mathscr{A})$; (III) допущение, что для $A\in\sigma(\mathscr{A})$ $\nu(A)<\mu^*(A)$ приводит к противоречию: $\mu^*(A)+\mu^*(\bar{A})>1$.

[При $A\in\mathscr{U}(\mathscr{A})$, т. е. $A=\bigcup_{(k)}A_k$, $A_1\subseteq A_2\subseteq\ldots$, $A_k\in\mathscr{A}$, $\mu^*(A)=\lim_{k\to\infty}\mu(A_k)$ по определению и $\nu(A)=\lim_{k\to\infty}\nu(A_k)$ по свойству непрерывности меры ν, что устанавливает утверждение (I). (II) При $A\in\sigma(\mathscr{A})$

$$\mu^*(A)=\inf\{\mu(B): A\subseteq B,\ \ B\in\mathscr{U}\}=$$

$$=\inf\{\nu(B): A\subseteq B,\ \ B\in\mathscr{U}\}\geqslant\nu(A).$$

(III) Предположив, что $\nu(A)<\mu^*(A)$ для некоторого $A\in\sigma(\mathscr{A})$, получаем

$$1=\nu(A)+\nu(\bar{A})<\mu^*(A)+\mu^*(\bar{A}),$$

что противоречит μ^*-измеримости множества A.]

С учетом 6.17 можно сформулировать следующую *теорему о существовании и единственности* продолжения счетно-аддитивной вероятностной меры μ, заданной на алгебре множеств $\mathscr{A}$ (или полуалгебре $\mathscr{D}$): на σ-алгебре $\sigma(\mathscr{A})$ существует единственная счетно-аддитивная мера, совпадающая на алгебре $\mathcal{A}$ с мерой μ; для этой меры на $\sigma(\mathcal{A})$ сохраняется обозначение μ. Этот результат верен и для любой конечной меры μ, так как от нее можно перейти к вероятностной мере $\mu(A)/\mu(\Omega)$.

Покажем, что теорема существования и единственности продолжения меры верна для *σ-конечной* счетно-аддитивной меры μ на $\mathcal{A}$ (или $\mathcal{D}$), т. е. такой меры, что пространство Ω можно разбить на счетное число непересекающихся элементов Ω_i из $\mathcal{A}$, для которых $\mu(\Omega_i)<\infty$, $i=1, 2, \ldots$. Обозначим $\mathcal{A}_i$ алгебру подмножеств из Ω_i, состоящую из всех $A\in\mathcal{A}$, $A\subseteq\Omega_i$. Мера μ, суженная на $\mathcal{A}_i$, конечна и потому допускает единственное продолжение на σ-алгебру $\sigma(\mathcal{A}_i)$ подмножеств из Ω_i. Рассмотрим класс $\mathcal{B}$ всех подмножеств из Ω, представимых в виде

$$B=\bigcup_{i=1}^{\infty} B_i, \quad B_i\in\sigma(\mathcal{A}_i), \quad i=1, 2, \ldots,$$

$\mathcal{B}$ содержит $\varnothing$, Ω. Если $B\in\mathcal{B}$, то

$$\bar{B}=\bigcup_{i=1}^{\infty}\Omega_i\bar{B}_i\in\mathcal{B};$$

если $B^k\in\mathcal{B}$, $k=1, 2, \ldots$, $B^kB^l=\varnothing$ при $k\neq l$, то

$$\bigcup_{k=1}^{\infty} B^k=\bigcup_{k=1}^{\infty}\bigcup_{i=1}^{\infty} B_i^k=\bigcup_{i=1}^{\infty}\Big(\bigcup_{k=1}^{\infty} B_i^k\Big)\in\mathcal{B}.$$

Следовательно, $\mathcal{B}$ есть σ-алгебра. Поскольку для любого $A\in\mathcal{A}$

$$A=\bigcup_{i=1}^{\infty} A\Omega_i\in\mathcal{B},$$

то $\mathcal{B}\supseteq\sigma(\mathcal{A})$. С другой стороны, $\sigma(\mathcal{A}_i)\subseteq\sigma(\mathcal{A})$, $i=1, 2, \ldots$, так что $\mathcal{B}\subseteq\sigma(\mathcal{A})$. Таким образом, $\mathcal{B}=\sigma(\mathcal{A})$. Положим для любого $B\in\mathcal{B}$

$$\hat{\mu}(B)=\sum_{i=1}^{\infty}\mu(B_i), \quad B_i\in\sigma(\mathcal{A}_i), \quad i=1, 2, \ldots.$$

Очевидно, функция $\hat{\mu}(B)$, $B\in\mathcal{B}$, — счетно-аддитивная мера на $\mathcal{B}$, совпадающая с μ на $\mathcal{A}$, и является искомым продолжением μ.

6.18. Задача. Доказать теорему единственности для σ-конечных мер.

[Пусть μ_1, μ_2 — σ-конечные меры на σ-алгебре множеств $\sigma(\mathcal{A})$, совпадающие на $\mathcal{A}$, $\Omega=\bigcup_{i=1}^{\infty}\Omega_i$, $\Omega_i\in\mathcal{A}$, $i=1, 2, \ldots$, $\Omega_i\Omega_j=$ $=\varnothing$ при $i\neq j$, $\mu_1(\Omega_i)=\mu_2(\Omega_i)<\infty$, $i=1, 2, \ldots$. В таком случае

$\mu_1=\mu_2$ на алгебре $\mathcal{A}_i=\{A : A\in\Omega_i,\ A\in\mathcal{A}\}$, а следовательно, $\mu_1=\mu_2$ на $\sigma(\mathcal{A}_i)$, $i=1, 2, \ldots$, Класс множеств вида

$$B=\bigcup_{i=1}^{\infty} B_i,\quad B_i\in\sigma(\mathcal{A}_i),\quad i=1, 2, \ldots,$$

совпадает с $\sigma(\mathcal{A})$ (см. выше), откуда

$$\mu_1(B)=\sum_{i=1}^{\infty}\mu_1(B_i)=\sum_{i=1}^{\infty}\mu_2(B_i)=\mu_2(B).]$$

6.19. Пример (мера Лебега в R^n). В 6.3, 6.4 было показано, что n-мерный объем — счетно-аддитивная мера на полуалгебре $\mathcal{D}_n$ параллелепипедов в R^n. Эта мера, очевидно, σ-конечна. Ее продолжение с $\mathcal{D}_n$ на $\sigma(\mathcal{D}_n)$ называют *мерой Лебега* в R^n. Класс множеств $\mathcal{B}^n=\sigma(\mathcal{D}_n)$, представляющий собой минимальную σ-алгебру, содержащую все n-мерные параллелепипеды, называют *борелевским* классом, а множества из $\sigma(\mathcal{D}_n)$ — *борелевскими* в честь французского математика Э. Бореля.

6.20. Задача. Пусть μ — счетно-аддитивная мера на классе борелевских множеств прямой, или полуалгебре промежутков, $\mu(I_{a,b})<\infty$ при $\infty<a<b<\infty$. Показать, что существует неубывающая непрерывная справа (слева) функция F на R, что для всех $a,\ b\in R$

$$\mu((a,\ b])=F(b)-F(a)\quad (\mu([a,\ b))=F(b)-F(a)).$$

Если μ конечна, то

$$\mu((-\infty,\ \infty))=F(\infty)-F(-\infty),\quad \mu((-\infty,\ b])=F(b)-F(-\infty),$$

где $F(\pm\infty)=\lim\limits_{x\to\pm\infty}F(x)$.

[Фиксируем произвольное $a\in R$ и определим $F(x)$ как

$$F(x)=\mu((a,\ x]),\quad x\geqslant a,\quad F(x)=-\mu((x,\ a]),\quad x\leqslant a.$$

Возьмем произвольную убывающую последовательность $x_n\to x>a$. Тогда $(a,\ x]=\bigcap\limits_{n=1}^{\infty}(a,\ x_n]$ и по непрерывности меры μ получаем $F(x_n)\to F(x)$, $n\to\infty$. Аналогично рассматриваются $x=a$ и $x<a$. Поскольку промежутки $(-n,\ n]$, $n=1, 2, \ldots$, расширяются к R, то по непрерывности

$$\mu(R)=\lim_{n\to\infty}\mu((-n,\ n])=\lim_{n\to\infty}(F(n)-F(-n))=F(+\infty)-F(-\infty).]$$

6.21*. Задача. Пусть F — неубывающая непрерывная справа (слева) действительная функция. Показать, что существует счетно-аддитивная мера μ, такая, что для любых $a,\ b\in R$

$$\mu((a,\ b])=F(b)-F(a)\quad (\mu([a,\ b))=F(b)-F(a)),$$

вводя полуалгебру промежутков вида $(-\infty, \infty)$, $(-\infty, b)$, $(a, b]$, (a, ∞) и действуя по схеме 6.3.

[Легко видеть, что μ есть σ-конечная конечно-аддитивная мера на полуалгебре, для применения теоремы о продолжении меры достаточно проверить ее счетную аддитивность на полуалгебре. Для конечного промежутка рассуждения полностью аналогичны 6.3, в случае бесконечного промежутка можно воспользоваться представлением

$$\mu(I) = \sum_{n=-\infty}^{\infty} \mu(I \cap (n, n+1]).]$$

6.22. Задача. Построить пример конечно-аддитивной меры на полуалгебре промежутков вида $(-\infty, \infty)$, $(-\infty, b]$, $(a, b]$ (a, ∞), не являющейся счетно-аддитивной.

6.23. Задача. Пусть μ — счетно-аддитивная мера на борелевских множествах прямой с конечными значениями на конечных промежутках. Показать, что имеется не более чем счетное число точек положительной меры: $\mu(\{x\}) > 0$.

6.24. Задача. Пусть Ω_i — некоторое множество, $\mathcal{A}_i$ — некоторая алгебра его подмножеств, $i=1, 2$. Положим

$$\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

и обозначим $\mathcal{A}_1 \times \mathcal{A}_2$ класс «прямоугольников» $A_1 \times A_2$, $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$. Показать, что $\mathcal{A}_1 \times \mathcal{A}_2$ образует полуалгебру.

[Если $A_1 \times A_2$, $A_1' \times A_2' \in \mathcal{A}_1 \times \mathcal{A}_2$, то (рис. 23)

$$\overline{A_1 \times A_2} = \{(\omega_1, \omega_2) : \omega_2 \overline{\in} A_2\} \cup \{(\omega_1, \omega_2) : \omega_2 \in A_2, \omega_1 \overline{\in} A_1\} =$$

$$= \{(\omega_1, \omega_2) : \omega_2 \in \overline{A}_2\} \cup \{(\omega_1, \omega_2) : \omega_1 \in \overline{A}_1, \omega_2 \in A_2\} =$$

$$= (\Omega_1 \times \overline{A}_2) \cup (\overline{A}_1 \times A_2); \quad (A_1 \times A_2) \cap (A_1' \times A_2') =$$

$$= \{(\omega_1, \omega_2) : \omega_1 \in A_1 \cap A_1', \quad \omega_2 \in A_2 \cap A_2'\} = (A_1 \cap A_1') \times (A_2 \cap A_2').]$$

6.25*. Задача. В условиях задачи 6.24 предположим, что на алгебре $\mathcal{A}$ подмножеств из Ω_i задана конечно-аддитивная вероятностная мера $\mathbf{P}_i$, $i=1, 2$. Определим на полуалгебре $\mathcal{A}_1 \times \mathcal{A}_2$ функцию $\mathbf{P} = \mathbf{P}_1 \times \mathbf{P}_2$:

$$\mathbf{P}(A_1 \times A_2) = \mathbf{P}_1(A_1)\mathbf{P}_2(A_2).$$

Показать, что $\mathbf{P}$ — конечно-аддитивная вероятностная мера на полуалгебре $\mathcal{A}_1 \times \mathcal{A}_2$.

[Пусть $A \times B = \bigcup_{i=1}^{n} A_i \times B_i$ — конечное разложение «прямоугольника» $A \times B \in \mathcal{A}_1 \times \mathcal{A}_2$ на непересекающиеся «прямоугольники» $A_i \times B_i \in \mathcal{A}_1 \times \mathcal{A}_2$, $i=1, 2, \ldots, n$. Надо показать, что

$$\mathbf{P}(A \times B) = \sum_{i=1}^{n} \mathbf{P}(A_i \times B_i).$$

Разберем вначале частный случай такого разложения. Допустим, что

$$A=\bigcup_{i=1}^{n} A_i, \quad A_iA_k=\varnothing, \quad i\neq k, \quad B=\bigcup_{j=1}^{m} B_j,$$

$$B_jB_l=\varnothing, \quad j\neq l, \quad A\times B=\bigcup_{(i,j)} A_i\times B_j.$$

Тогда

$$\sum_{(i,j)} \mathbf{P}(A_i\times B_j)=\sum_{i=1}^{n}\sum_{j=1}^{m}\mathbf{P}(A_i)\,\mathbf{P}(B_j)=\mathbf{P}(A)\,\mathbf{P}(B).$$

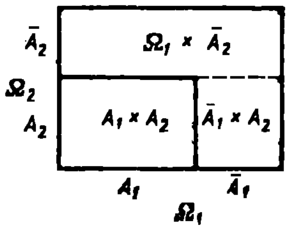

Рис. 23

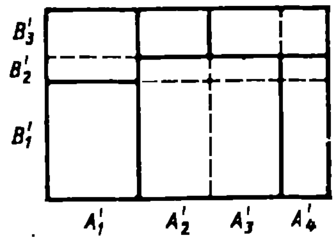

Рис. 24

Общий случай сведем к разобранному, рассматривая разбиения

$$\bigcup_{(\delta_1,\ldots,\delta_n)} A_1^{\delta_1}\ldots A_n^{\delta_n}, \quad \bigcup_{(\delta_1,\ldots,\delta_n)} B_1^{\delta_1}\ldots B_n^{\delta_n},$$

где объединение берется по $\delta_1, \ldots, \delta_n$, принимающим значения 0 и 1, и где мы полагаем $A^0=A$, $A^1=\bar{A}$. Множества $A_1^{\delta_1}\ldots A_n^{\delta_n}$ при различных наборах $(\delta_1, \delta_2, \ldots, \delta_n)$ не пересекаются, каждое A_i можно представить в виде

$$A_i=\bigcup_{(\delta_1,\ldots,\delta_{i-1},\delta_{i+1},\ldots,\delta_n)} A_1^{\delta_1}\ldots A_{i-1}^{\delta_{i-1}}A_iA_{i+1}^{\delta_{i+1}}\ldots A_n^{\delta_n}.$$

В самом деле, левая часть, очевидно, содержит правую. С другой стороны, если $\omega_1\in A_i$, то для любого $j\neq i$ либо $\omega_1\in A_j$, либо $\omega_1\in \bar{A}_j$, так что если определить δ'_j условием $\omega_1\in A_j^{\delta'_j}$, то ω_1 входит в правую часть в член с индексами $\delta'_1, \ldots, \delta'_{i-1}, \delta'_{i+1}, \ldots, \delta'_n$. Введем обозначения $A'_1, \ldots, A'_{n'}$ для непустых множеств $A_1^{\delta_1}\ldots A_n^{\delta_n}$ и $B'_1, \ldots, B'_{m'}$ для непустых множеств $B_1^{\delta_1}\ldots B_n^{\delta_n}$, где наборы $(\delta_1, \ldots, \delta_n)$ пробегают все 2^n возможных значений. Таким образом, множества A'_k, $k=1, \ldots, n'$, попарно не пересекаются, это же верно для множеств B'_l, $l=1, \ldots, m'$, и при этом для любого i (рис. 24)

$$A_i=\bigcup_{(k\in K(i))} A'_k, \quad B_i=\bigcup_{(l\in L(i))} B'_l,$$

где $K(i) \subseteq \{1, 2, \ldots, n'\}$, $L(i) \subseteq \{1, 2, \ldots, m'\}$ — некоторые множества индексов. Поскольку

$$A = \bigcup_{i=1}^{n} A_i \subseteq \bigcup_{k=1}^{n'} A'_k, \quad B = \bigcup_{i=1}^{n} B_i \subseteq \bigcup_{l=1}^{m'} B'_l,$$

то

$$A \times B \subseteq \bigcup_{k=1}^{n'} \bigcup_{l=1}^{m'} A'_k \times B'_l.$$

Обратное включение очевидно, так что имеем разбиение $A \times B$ на «клетки» (рис. 24):

$$A \times B = \bigcup_{k=1}^{n'} \bigcup_{l=1}^{m'} A'_k \times B'_l.$$

Точно так же разбиваются на «клетки» множества

$$A_i \times B_i = \bigcup_{\{(k,l) \in K(i) \times L(i)\}} A'_k \times B'_l, \quad i = 1, 2, \ldots, n.$$

Так как множества $A_i \times B_i$, $i=1, 2, \ldots, n$, попарно не пересекаются, то и множества $K(i) \times L(i)$, $i=1, 2, \ldots, n$, попарно не пересекаются, а из равенства $A \times B = \bigcup\limits_{i=1}^{n} A_i \times B_i$ вытекает, что

$$\{1, 2, \ldots, n'\} \times \{1, 2, \ldots, m'\} = \bigcup_{i=1}^{n} K(i) \times L(i).$$

Используя разобранный ранее частный случай, получаем

$$\mathbf{P}(A \times B) = \sum_{k=1}^{n'} \sum_{l=1}^{m'} \mathbf{P}(A'_k \times B'_l) =$$

$$= \sum_{i=1}^{n} \sum_{\{(k,l) \in K(i) \times L(i)\}} \mathbf{P}(A'_k \times B'_l) = \sum_{i=1}^{n} \mathbf{P}(A_i \times B_i).]$$

6.26. Пример (независимые испытания с произвольным множеством исходов). Построим модель независимых испытаний, описываемых последовательностью вероятностных пространств $(\Omega_i, \mathscr{B}_i, \mathbf{P}_i)$, $i=1, 2, \ldots, n$, общего вида. Ради упрощения обозначений рассмотрим $n=2$, общий случай легко получается по индукции. Определим на полуалгебре $\mathscr{B}_1 \times \mathscr{B}_2$ подмножеств из $\Omega_1 \times \Omega_2$ конечно-аддитивную меру $\mathbf{P} = \mathbf{P}_1 \times \mathbf{P}_2$, как это сделано в 6.25. Продолжим меру $\mathbf{P}$ на алгебру $\mathscr{A}$, порожденную полуалгеброй $\mathscr{B}_1 \times \mathscr{B}_2$, и докажем непрерывность меры $\mathbf{P}$ в $\varnothing$.

Докажем *ключевое неравенство*. Пусть $A \in \mathscr{A}$, для любого $\omega_1 \in \Omega_1$ введем *сечение множества A координатой* ω_1 (рис. 25).

$$A(\omega_1) = \{\omega_2 : (\omega_1, \omega_2) \in A\}$$

и определим *проекцию множества* A *на* Ω_1 как

$$\pi_1(A)=\{\omega_1 : A(\omega_1)\neq\varnothing\}\subseteq\Omega_1.$$

Покажем, что если для всех $\omega_1\in\Omega_1$ имеет место неравенство $\mathbf{P}_2(A(\omega_1))\leqslant\alpha$, то для вероятности $\mathbf{P}(A)$ справедлива оценка

$$\mathbf{P}(A)\leqslant\mathbf{P}_1(\pi_1(A))\cdot\alpha.$$

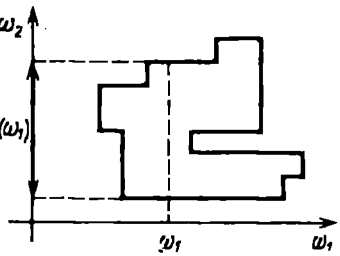

Рис. 25

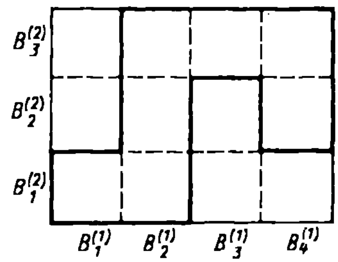

Рис. 26

Нетрудно проверить, что A можно представить в виде (рис. 26)

$$A=\bigcup_{\{(i,j)\in M\}} B_i^{(1)}\times B_j^{(2)},$$

где $B_i^{(1)}\in\mathscr{B}_1$, $i=1,\ldots,n$, $B_i^{(1)}B_j^{(1)}=\varnothing$ при $i\neq j$, $B_j^{(2)}\in\mathscr{B}_2$, $j=1,\ldots,m$, $B_j^{(2)}B_l^{(2)}=\varnothing$ при $j\neq l$ (для доказательства применима конструкция разбиения на клетки из 6.25). Отсюда получаем

$$\mathbf{P}(A)=\sum_{\{(i,j)\in M\}}\mathbf{P}_1(B_i^{(1)})\cdot\mathbf{P}_2(B_j^{(2)})=$$

$$=\sum_{\{i\,:\,B_i^{(1)}\subseteq\pi_1(A)\}}\mathbf{P}_1(B_i^{(1)})\sum_{\{j\,:\,(i,j)\in M\}}\mathbf{P}_2(B_j^{(2)}).$$

С другой стороны, поскольку ω_1-сечение любого прямоугольника $B_1\times B_2$ либо пусто, либо есть B_2, то при $\omega_1\in B_i^{(1)}$ получаем

$$A(\omega_1)=\bigcup_{\{j\,:\,(i,j)\in M\}} B_j^{(2)},$$

откуда

$$\alpha\geqslant\mathbf{P}_2(A(\omega_1))=\sum_{\{j\,:\,(i,j)\in M\}}\mathbf{P}_2(B_j^{(2)}),\quad \omega_1\in B_i^{(1)}.$$

Подставляя полученное неравенство в выражение для $P(A)$, имеем

$$\mathbf{P}(A)\leqslant\alpha\sum_{\{i\,:\,B_i^{(1)}\in\pi_1(A)\}}\mathbf{P}_1(B_i^{(1)})=\alpha\mathbf{P}_1(\pi_1(A)).$$

Пусть дано некоторое $\varepsilon>0$, положим

$$\hat{A}=\{\omega_1 : \mathbf{P}_2(A(\omega_1))>\varepsilon/2\}\subseteq\Omega_1.$$

Легко видеть что $\hat{A}=\mathscr{B}_1$ (см. выше представление A в виде объединения прямоугольников). Так как

$$A=(A\cap(\hat{A}\times\Omega_2))\cup(A\cap(\bar{\hat{A}}\times\Omega_2)),$$

то

$$\mathbf{P}(A)=\mathbf{P}(A\cap(\hat{A}\times\Omega_2))+\mathbf{P}(A\cap(\bar{\hat{A}}\times\Omega_2))\leqslant\mathbf{P}_1(\hat{A})+\mathbf{P}_1(\bar{\hat{A}})\cdot\varepsilon/2,$$

где мы воспользовались выведенным выше неравенством в применении к множеству $A\cap(\bar{\hat{A}}\times\Omega_2)$. Отсюда получаем *ключевое неравенство*

$$\mathbf{P}_1(\hat{A})\geqslant\mathbf{P}(A)-\varepsilon/2.$$

Возьмем произвольную убывающую последовательность $A_i\in\mathscr{A}$, $i=1, 2, \ldots, \bigcap_{i=1}^{\infty}A_i=\varnothing$. Предположим, что для некоторого $\varepsilon>0$ $\mathbf{P}(A_i)>\varepsilon>0$ при всех i, и приведем это предположение к противоречию, показав, что множества A_i, $i=1, 2, \ldots,$ имеют общую точку. Для каждого A_i, $i=1, 2, \ldots$, введем множество

$$\hat{A}_i=\{\omega_1 : \mathbf{P}_2(A_i(\omega_1))>\varepsilon/2\}.$$

Поскольку последовательность A_i не убывает, то при любом ω_1 не убывает последовательность $A_i(\omega_1)$, а значит, не убывает и последовательность $\hat{A}_i$, $i=1, 2, \ldots$. Применяя к A_i, $\hat{A}_i$ установленное выше неравенство, имеем

$$\mathbf{P}_1(\hat{A}_i)\geqslant\mathbf{P}(A_i)-\varepsilon/2>\varepsilon/2.$$

В таком случае для множества $\hat{A}=\bigcap_{i=1}^{\infty}\hat{A}_i$ имеем $\mathbf{P}_1(\hat{A})=\lim_{i\to\infty}\mathbf{P}(\hat{A}_i)\geqslant\geqslant\varepsilon/2$ и, в частности, $\hat{A}$ не пусто. Обозначим $\tilde{\omega}_1$ общую точку множеств $\hat{A}_i$, $i=1, 2, \ldots$. В силу определения множеств $\hat{A}_i$ имеем

$$\mathbf{P}_2(A_i(\tilde{\omega}_1))\geqslant\varepsilon/2,\quad i=1, 2, \ldots.$$

Учитывая, что последовательность $A_i(\tilde{\omega}_1)$, $i=1, 2, \ldots$, не убывает, получаем

$$\mathbf{P}_2\left(\bigcap_{i=1}^{\infty}A_i(\tilde{\omega})\right)\geqslant\varepsilon/2.$$

Следовательно, $\bigcap_{i=1}^{\infty}A_i(\tilde{\omega}_1)\neq\varnothing$. Пусть $\tilde{\omega}_2$ — общая точка всех множеств $A_i(\tilde{\omega}_1)$, $i=1, 2, \ldots$. Так как $\tilde{\omega}_1\times A_i(\tilde{\omega}_1)\subseteq A_i$, $i=1, 2, \ldots$, то $(\widetilde{\omega}_1, \widetilde{\omega}_2) \in A_i$, $i=1, 2, \ldots$. Полученное противоречие устанавливает непрерывность в $\varnothing$ меры $\mathbf{P}$ на алгебре $\mathscr{A}$. Применяя теорему о продолжении меры, получаем вероятностное пространство $(\Omega_1 \times \Omega_2, \sigma(\mathscr{B}_1 \times \mathscr{B}_2), \mathbf{P})$.

6.27. Пример (бесконечная последовательность испытаний Бернулли). Положим $\mathscr{X} = \{0, 1\}$, $\mathscr{X}^n = \mathscr{X} \times \ldots \times \mathscr{X}$ (n раз), $\mathscr{X}^\infty = \mathscr{X} \times \times \mathscr{X} \times \ldots$. Пусть $(\mathscr{X}^n, \mathbf{P}^{(n)})$ есть вероятностное пространство n испытаний Бернулли:

$$\mathbf{P}^{(n)}(B) = \sum_{(x_1, \ldots, x_n) \in B} p(x_1, \ldots, x_n), \quad B \subseteq \mathscr{X}^n,$$

$$p(x_1, \ldots, x_n) = p^{s_n}(1-p)^{n-s_n}, \quad s_n = x_1 + \ldots + x_n.$$

Введем алгебру $\mathscr{A}$ подмножеств $A \subseteq \mathscr{X}^\infty$ вида (см. 2.15, 2.16)

$$A = \{\mathbf{x} = (x_1, x_2, \ldots) : (x_1, \ldots, x_n) \in A^{(n)}\}, \quad A^{(n)} \subseteq \mathscr{X}^n,$$

и определим конечно-аддитивную вероятностную меру $\mathbf{P}(A) = \mathbf{P}^{(n)}(A^{(n)})$. Образуем сечение

$$A^{(n)}(x_1) = \{(x_2, \ldots, x_n) : (x_1, \ldots, x_n) \in A^{(n)}\}$$

множества $A^{(n)} \subseteq \mathscr{X}^n$ координатой x_1 и положим

$$\widehat{A}^{(n)} = \{x_1 : \mathbf{P}^{(n-1)}(A^{(n)}(x_1)) > \varepsilon/2\},$$

где $\varepsilon > 0$ фиксировано. Выведем ключевое неравенство (ср. 6.26):

$$\mathbf{P}^{(n)}(A^{(n)}) = \sum_{\{x_1\}} p(x_1) . \mathbf{P}^{(n-1)}(A^{(n)}(x_1)) \leqslant \mathbf{P}^{(1)}(\widehat{A}^{(n)}) +$$

$$+ (1 - \mathbf{P}^{(1)}(\widehat{A}^{(n)})) . \varepsilon/2; \quad \mathbf{P}^{(1)}(\widehat{A}^{(n)}) \geqslant \mathbf{P}^{(n)}(A^{(n)}) - \varepsilon/2.$$

Чтобы воспользоваться теоремой о продолжении меры $\mathbf{P}$ с алгебры $\mathscr{A}$ на $\sigma(\mathscr{A})$, достаточно проверить, что мера P непрерывна в $\varnothing$, т. е. для любой последовательности

$$A_1 \supseteq A_2 \supseteq \ldots, \quad A_i \in \mathscr{A}, \quad i = 1, 2, \ldots, \quad \bigcap_{i=1}^{\infty} A_i = \varnothing$$

имеем $\mathbf{P}(A_i) \to 0$ при $i \to \infty$. По определению алгебры $\mathscr{A}$ каждое A_i можно представить в виде

$$A_i = \{\mathbf{x} = (x_1, x_2, \ldots) : (x_1, \ldots, x_{n_i}) \in A^{(n_i)}\}, \quad A^{(n_i)} \subseteq \mathscr{X}^{n_i}.$$

Покажем, что без ограничения общности можно считать $n_i = i$. Во-первых, легко видеть, что последовательность n_i, $i = 1, 2, \ldots$, можно взять возрастающей (ср. 2.16). Предполагая, что $n_1 <$

$<n_2<\ldots$, образуем последовательность $\widetilde{A}_1 \supseteq \widetilde{A}_2 \supseteq \ldots$, полагая $\widetilde{A}_i$ равным i-му члену последовательности

$$\underbrace{\mathscr{X}, \ldots, \mathscr{X}}_{n_1-1}, A_1, \underbrace{A_1, \ldots, A_1}_{n_2-n_1-1}, A_2, \underbrace{A_2, \ldots, A_2}_{n_3-n_2-1}, A_3, \ldots .$$

Последовательность $\widetilde{A}_i$, $i=1, 2, \ldots$, такова, что в представлении

$$\widetilde{A}_i = \{\mathbf{x} : (x_1, \ldots, x_{n_i}) \in \widetilde{A}^{(n_i)}\}, \quad \widetilde{A}^{(n_i)} \subseteq \mathscr{X}^{n_i},$$

можно взять $n_i = i$. Так как

$$\bigcap_{i=1}^{\infty} A_i = \bigcap_{i=1}^{\infty} \widetilde{A}_i, \quad \lim_{i\to\infty} \mathbf{P}(A_i) = \lim_{i\to\infty} \mathbf{P}(\widetilde{A}_i),$$

то достаточно установить, что $\mathbf{P}(\widetilde{A}_i) \to 0$ при $i\to\infty$, т. е. исходную последовательность можно заменить на последовательность с требуемым свойством. Чтобы упростить обозначения, положим $\widetilde{A}_i = A_i$.

Предположим, что $\mathbf{P}(A_n) \not\to 0$ при $n\to\infty$, т. е. $\mathbf{P}(A_n) = \mathbf{P}^{(n)}(A^{(n)}) > \varepsilon$ при некотором $\varepsilon>0$ и всех n, и приведем это к противоречию, построив точку $\widetilde{\mathbf{x}} = (\widetilde{x}_1, \widetilde{x}_2, \ldots)$, общую для всех A_n. Применяя ключевое неравенство, имеем

$$\mathbf{P}^{(1)}(\widehat{A}_n^{(n)}) \geqslant \mathbf{P}^{(n)}(A_n^{(n)}) - \varepsilon/2 > \varepsilon/2.$$

Так как $A_n \supseteq A_{n+1}$, то $A_n^{(n)} \times \{0, 1\} \supseteq A_{n+1}^{(n+1)}$ и, следовательно, этим же знаком включения связаны их сечения:

$$A_n^{(n)}(x_1) \times \{0, 1\} \supseteq A_{n+1}^{(n+1)}(x_1).$$

Поэтому

$$\mathbf{P}^{(n-1)}(A_n^{(n)}(x_1)) \geqslant \mathbf{P}^{(n)}(A_{n+1}^{(n+1)}(x_1)).$$

Отсюда получаем, что

$$\widehat{A}_{n+1}^{(n+1)} \subseteq \widehat{A}_n^{(n)}.$$

В таком случае

$$\mathbf{P}^{(1)}\Bigl(\bigcap_{n=1}^{\infty} \widehat{A}_n^{(n)}\Bigr) \geqslant \varepsilon/2,$$

и существует точка $\widetilde{x}_1$, общая для всех множеств $\widehat{A}_n^{(n)}$.

Последовательность сечений

$$A_n(\widetilde{x}_1) = \{(x_2, x_3, \ldots) : (\widetilde{x}_1, x_2, \ldots) \in A_n\}$$

— убывающая. По определению множества $\widehat{A}_n^{(n)}$ имеем

$$\mathbf{P}(A_n(\widetilde{x}_1)) \equiv \mathbf{P}^{(n-1)}(A_n^{(n)}(\widetilde{x}_1)) > \varepsilon/2.$$

Кроме того, $\tilde{x}_1 \times A_n(\tilde{x}_1) \subseteq A_n$. Проводя для последовательности $B_n \equiv \equiv A_n(\tilde{x}_1)$, $n=1, 2, \ldots$, те же рассуждения, что и для A_n, $n = = 1, 2, \ldots$, построим общую для всех $\widehat{B}_n^{(n)}$ точку $\tilde{x}_2$. При этом последовательность множеств

$$B_n(\tilde{x}_2) = \{(x_3, x_4, \ldots) : (\tilde{x}_2, x_3, \ldots) \in B_n\}$$

убывает и

$$\mathbf{P}(B_n(\tilde{x}_2)) > \varepsilon/2^2, \quad \tilde{x}_2 \times B_n(\tilde{x}_2) \subseteq B_n, \quad n = 1, 2, \ldots .$$

Поскольку $A_n(\tilde{x}_1, \tilde{x}_2) = B_n(\tilde{x}_2)$, то $\tilde{x}_1 \times \tilde{x}_2 \times A_n(\tilde{x}_1, \tilde{x}_2) \subseteq A_n$. Продолжая этот процесс до бесконечности, получим точку $(\tilde{x}_1, \tilde{x}_2, \ldots)$, общую для всех A_n, $n=1, 2, \ldots$.

6.28. Задача. Построить модель бесконечной последовательности независимых испытаний с произвольным множеством исходов.

[Выводится из 6.26 по схеме 6.27.]

6.29 Пример (бесконечная последовательность зависимых испытаний с дискретным множеством исходов). Пусть $\mathscr{X}$ — конечное или счетное множество, $\mathscr{X}^n = \mathscr{X} \times \ldots \times \mathscr{X}$ (n раз), $\mathscr{X}^\infty = = \mathscr{X} \times \mathscr{X} \times \ldots$. Предположим, что имеется семейство дискретных вероятностных пространств $(\mathscr{X}^n, p^{(n)}(x_1, \ldots, x_n))$, причем выполняются так называемые *условия согласованности*:

$$\sum_{\{x_{n+1}\}} p^{(n+1)}(x_1, \ldots, x_{n+1}) = p^{(n)}(x_1, \ldots, x_n), \quad n = 1, 2, \ldots,$$

которые означают, что при любых $n<m$ пространство $(\mathscr{X}^n, p^{(n)})$ может служить описанием первых n испытаний опыта, представленного пространством $(\mathscr{X}^m, p^{(m)})$. На алгебре $\mathscr{A}$ подмножеств $A \subseteq R^\infty$, имеющих вид

$$A = \{\mathbf{x} = (x_1, x_2, \ldots) : (x_1, \ldots, x_n) \in A^{(n)}\}, \quad A^{(n)} \subseteq \mathscr{X}^n,$$

определим

$$\mathbf{P}(A) = \mathbf{P}^{(n)}(A^{(n)}) = \sum_{(x_1, \ldots, x_n) \in A^{(n)}} p^{(n)}(x_1, \ldots x_n).$$

Легко проверить, что мера $\mathbf{P}$ определена корректно и конечно-аддитивна на $\mathscr{A}$. Чтобы продолжить ее на $\sigma(\mathscr{A})$, действуем по уже известной схеме. Введем на $\mathscr{X}^{n-1}$ семейство условных вероятностных мер

$$\mathbf{P}^{(n-1)}(A^{(n-1)} \mid x_1) = \sum_{(x_2, \ldots, x_n) \in A^{(n-1)}} p^{(n)}(x_1, \ldots, x_n)/p^{(1)}(x_1)$$

при $x_1 \in \mathscr{X}$, таких, что $p^{(1)}(x_1) > 0$. Далее запишем

$$A^{(n)}(x_1) = \{(x_2, \ldots, x_n) : (x_1, \ldots, x_n) \in A^{(n)}\},$$

$$\widehat{A}^{(n)} = \{x_1 : \mathbf{P}^{(n-1)}(A^{(n)}(x_1) \mid x_1) > \varepsilon/2\},$$

$$\mathbf{P}^{(n)}(A^{(n)}) = \sum_{\{x_1\}} p^{(1)}(x_1) \cdot \mathbf{P}^{(n-1)}(A^{(n)}(x_1) \mid x_1) \leqslant$$

$$\leqslant \mathbf{P}^{(1)}(\widehat{A}^{(n)}) + (1 - \mathbf{P}^{(1)}(\widehat{A}^{(n)})) \cdot \varepsilon/2.$$

Дальнейшие рассуждения аналогичны проведенным в 6.27.

6.30. Задача. Класс множеств $\mathscr{M}$ называется *монотонным*, если он замкнут относительно счетных объединений возрастающих последовательностей и счетных пересечений убывающих последовательностей:

$$A_i \in \mathscr{M}, \quad A_i \subseteq A_{i+1}, \quad i = 1, 2, \ldots \to \bigcup_{\{i\}} A_i \in \mathscr{M},$$

$$B_i \in \mathscr{M}, \quad B_i \supseteq B_{i+1}, \quad i = 1, 2, \ldots \to \bigcap_{\{i\}} B_i \in \mathscr{M}.$$

Показать, что для любой системы S подмножеств из Ω существует наименьший содержащий его монотонный класс, который обозначим $\mathscr{M}(S)$.

[Рассуждения аналогичны проведенным при доказательстве существования $\sigma(S)$ (см. 6.16 и далее).]

6.31. Задача. Пусть $\mathscr{M}$ — некоторый монотонный класс подмножеств из Ω. Для произвольного $A \in \mathscr{M}$ положим

$$\mathscr{M}_A = \{B : B,\ BA,\ B\bar{A},\ \bar{B}A \in \mathscr{M}\}.$$

Показать, что $\mathscr{M}_A$ — монотонный класс.

[Пусть $B_i \in \mathscr{M}_A$, $B_i \subseteq B_{i+1}$, $i = 1, 2, \ldots$, так что

$$B_i,\ B_iA,\ B_i\bar{A},\ \bar{B}_iA \in \mathscr{M},\quad B_iA \subseteq B_{i+1}A,$$

$$B_i\bar{A} \subseteq B_{i+1}\bar{A},\quad \bar{B}_iA \supseteq \bar{B}_{i+1}A,\quad i = 1, 2, \ldots .$$

Тогда $B = \bigcup_{\{i\}} B_i \in \mathscr{M}_A$, так как

$$B = \bigcup_{\{i\}} B_i \in \mathscr{M},\quad BA = \bigcup_{\{i\}} B_iA \in \mathscr{M},\quad B\bar{A} = \bigcup_{\{i\}} B_i\bar{A} \in \mathscr{M},$$

$$\bar{B}A = \overline{\bigcup_{\{i\}} B_i}A = \bigcap_{\{i\}} \bar{B}_iA \in \mathscr{M}.$$

Точно так же убеждаемся, что класс $\mathscr{M}_A$ замкнут относительно пересечений убывающих последовательностей.]

6.32. Задача. Пусть $\mathscr{M} = \mathscr{M}(\mathscr{A})$ — минимальный монотонный класс, содержащий алгебру множеств $\mathscr{A}$. Показать, что $\mathscr{M}_A = \mathscr{M}(\mathscr{A})$ для любого $A \in \mathscr{A}$ (где класс $\mathscr{M}_A$ определен в 6.31).

[Пусть $A \in \mathscr{A}$ фиксировано. Так как для любого $B \in \mathscr{A}$ имеем BA, $B\bar{A}$, $\bar{B}A \in \mathscr{A} \subseteq \mathscr{M}(\mathscr{A})$, то $\mathscr{A} \subseteq \mathscr{M}_A$, где $\mathscr{M}_A$ ввиду 6.31 — монотонный класс. Но тогда $\mathscr{M}(\mathscr{A}) \subseteq \mathscr{M}_A$. Обратное включение следует из определения класса $\mathscr{M}_A$.]

6.33. Задача. Заметив, что для A, $B \in \mathscr{M}(\mathscr{A})$

$$A \in \mathscr{M}_B \Leftrightarrow B \in \mathscr{M}_A,$$

показать, используя 6.32, что $\mathscr{M}_B = \mathscr{M}_A$.

[Достаточно проверить, что $\mathscr{M}_B \supseteq \mathscr{M}(\mathscr{A})$, так как обратное включение следует из определения $\mathscr{M}_B$. Для любого $A \in \mathscr{M}(\mathscr{A})$ и любого $B \in \mathscr{M}(\mathscr{A}) = \mathscr{M}_A$ имеем $A \in \mathscr{M}_B$, т. е. $\mathscr{M}(\mathscr{A}) \subseteq \mathscr{M}_B$.]

6.34. Задача. Если $\mathscr{A}$ — алгебра множеств, то $\mathscr{M}(\mathscr{A})$ — также алгебра множеств.

[Для любых A_1, $A_2 \in \mathscr{M}(\mathscr{A})$, используя равенство $\mathscr{M}(\mathscr{A}) = \mathscr{M}_{A_1}$ (см. 6.33), имеем $A_2 \in \mathscr{M}_{A_1}$, т. е. A_1A_2, $A_1\bar{A}_2$, $\bar{A}_1A_2 \in \mathscr{M}(\mathscr{A})$. В частности, при $A_2 = \Omega$ получаем отсюда, что $\bar{A}_1 \in \mathscr{M}(\mathscr{A})$, так что класс $\mathscr{M}(\mathscr{A})$ замкнут относительно пересечений и перехода к дополнению и, следовательно, является алгеброй множеств.]

6.35. Задача. Если алгебра множеств $\mathscr{A}$ — монотонный класс, то она является σ-алгеброй.

[Пусть $A_i \in \mathscr{A}$, $i = 1, 2, \ldots$, покажем, что $A = \bigcup_{\{i\}} A_i \in \mathscr{A}$. Образуем последовательность попарно непересекающихся множеств $B_i = A_i\bar{A}_{i-1} \ldots \bar{A}_1 \in \mathscr{A}$, $i = 1, 2, \ldots$. Очевидно, что $\bigcup_{\{i\}} B_i = A$, кроме того, последовательность $C_n = \bigcup_{\{i \leqslant n\}} B_i \in \mathscr{A}$, возрастает и $\bigcup_{\{n\}} C_n = A$. Поскольку $\mathscr{A}$ — монотонный класс, то отсюда следует, что $A \in \mathscr{A}$.]

6.36. Задача. Если $\mathscr{A}$ — алгебра множеств, то $\mathscr{M}(\mathscr{A}) = \sigma(\mathscr{A})$.

[Из 6.34, 6.35 вытекает, что $\mathscr{M}(\mathscr{A})$ является σ-алгеброй множеств (содержащей $\mathscr{A}$), так что $\mathscr{M}(\mathscr{A}) \supseteq \sigma(\mathscr{A})$. Обратное включение следует из того, что любая σ-алгебра — монотонный класс.]

6.37. Задача. Пусть μ_1 и μ_2 — вероятностные счетно-аддитивные меры на $\sigma(\mathscr{A})$, где $\mathscr{A}$ — некоторая алгебра множеств, такие, что $\mu_1(A) = \mu_2(A)$ для всех $A \in \mathscr{A}$. Показать, что $\mu_1 = \mu_2$ на $\sigma(\mathscr{A})$.

[Обозначим $\mathscr{B}$ класс всех множеств $B \in \sigma(\mathscr{A})$, для которых $\mu_1(B) = \mu_2(B)$. Из свойства непрерывности мер μ_1, μ_2 вытекает, что $\mathscr{B}$ — монотонный класс. Поскольку $\mathscr{A} \subseteq \mathscr{B}$, то $\mathscr{B} \supseteq \mathscr{M}(\mathscr{A}) = \sigma(\mathscr{A})$.]

Отметим, что из 6.37 вытекает теорема о единственности продолжения счетно-аддитивной меры с алгебры множеств на минимальную σ-алгебру, доказанная другим способом в 6.17.

Глава II

ХАРАКТЕРИСТИКИ ВЕРОЯТНОСТНЫХ РАСПРЕДЕЛЕНИЙ

§ 7. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ

Распределение вероятностей случайной величины представляет собой достаточно сложный для описания и статистической оценки объект. Это делает желательным введение каких-либо интегральных характеристик, вбирающих в себя наиболее существенные для решения тех или иных задач черты распределения. К их числу относится *математическое ожидание* или *среднее значение* сл. в. $X(\omega)$. Пусть для начала сл. в. $X(\omega)$ определена на вероятностном пространстве с конечным числом равновозможных исходов. Среднее арифметическое значение $X(\omega)$ по всем N точкам пространства Ω называется *средним значением* или *математическим ожиданием* сл. в. X и обозначается

$$\mathbf{M}X = N^{-1}\sum_{\{\omega\in\Omega\}} X(\omega) = \sum_{\{\omega\in\Omega\}} X(\omega)\,p(\omega),$$

где $p(\omega) = N^{-1}$, $\omega\in\Omega$, — равномерное распределение вероятностей на Ω.

7.1. Задача. Сл. в. X — число красных шаров в n-кратной выборке с возвращением из урны с долей p красных шаров. Найти $\mathbf{M}X$.

[За пространство элементарных событий примем множество всех последовательностей $\omega = (i_1, \ldots, i_n)$ из номеров извлекаемых шаров, так что $p(\omega) = N^{-n}$ и (см. 1.21)

$$\mathbf{M}X = \sum_{\{\omega\}} X(\omega)\,p(\omega) = \sum_{k=0}^{n} k \sum_{\{\omega\,:\,X(\omega)=k\}} p(\omega) =$$

$$= \sum_{k=0}^{n} kp(\omega : X(\omega) = k) = \sum_{k=1}^{n} kC_n^k p^k (1-p)^{n-k} =$$

$$= \sum_{k=1}^{n} npC_{n-1}^{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} = np.]$$

Пусть $(\Omega, p(\omega))$ — произвольное дискретное вероятностное пространство, $X(\omega)$ — случайная величина. В предположении, что

$$\sum_{\{\omega\in\Omega\}} |X(\omega)|p(\omega) < \infty,$$

определим *математическое ожидание* сл. в. X формулой

$$MX = \sum_{\{\omega\in\Omega\}} X(\omega)p(\omega).$$

Как и в 7.1, получаем следующую формулу для вычисления математического ожидания:

$$MX = \sum_{\{x\in\mathscr{X}\}} \sum_{\{\omega : X(\omega=x)\}} X(\omega)p(\omega) = \sum_{\{x\in\mathscr{X}\}} xp_X(x).$$

Если $p(x)$ — произвольное дискретное распределение вероятностей, заданное на некотором подмножестве $\mathscr{X}\subseteq R$, то число

$$\sum_{\{x\in\mathscr{X}\}} xp(x)$$

в предположении, что ряд абсолютно сходится, называется *средним значением распределения* $p(x)$. Отметим, что математическое ожидание сл. в. X есть в то же время среднее значение ее распределения вероятностей $p_X(x)$.

7.2. Задача. Вычислить средние значения: (I) гипергеометрического распределения $p(k)=C_L^k C_{N-L}^{n-k}/C_N^n$, $k=0, 1, 2, \ldots$; (II) отрицательного биномиального распределения $p(k)=C_{r+k-1}^k p^r(1-p)^k$, $k=0, 1, 2, \ldots$; (III) пуассоновского распределения $p(k)=\lambda^k e^{-\lambda}/k!$, $k=0, 1, 2, \ldots$.

[(I) $\sum kC_L^k C_{N-L}^{n-k}/C_N^n = L\sum C_{L-1}^{k-1} C_{(N-1)-(L-1)}^{(n-1)-(k-1)} (C_{N-1}^{n-1}\cdot N/n)^{-1} = nL/N$;

(II) $$\sum_{k=0}^{\infty} kC_{r+k-1}^k p^r(1-p)^k = \sum_{k=1}^{\infty} \frac{(r+k-1)!}{(k-1)!\,r!} p^{r+1}(1-p)^{k-1)} \times$$
$$\times p^{-1}(1-p)r = p^{-1}(1-p)r;$$

(III) $$\sum_{k=0}^{\infty} k\lambda^k e^{-\lambda}/k! = \lambda\sum_{k=1}^{\infty} \lambda^{k-1}e^{-\lambda}/(k-1)! = \lambda.]$$

Роль математического ожидания как характеристики распределения вероятностей в немалой степени связана с тем, что, как и вероятность, это понятие имеет статистическое толкование. Заметим прежде всего, что если сл. в. X — индикаторная, то $MX=p=$ $=\mathbf{P}(\omega : X(\omega)=1)$, так что эмпирический закон устойчивости частот можно записать в виде $n^{-1}(x_1+\ldots+x_n)\approx MX$, где $x_1, \ldots, x_n$ —

последовательность повторных независимых наблюдений за сл. в. X.

Обобщая, рассмотрим дискретную сл. в. X, принимающую конечный набор значений a_j, $j=1, \ldots, m$, с вероятностями p_j, $j=1, \ldots, m$. Обозначив $\nu_j^{(n)}$ частоту исхода a_j в последовательности наблюдений $x_1, x_2, \ldots, x_n$ и записав равенство

$$n^{-1}(x_1+\ldots+x_n)=a_1\nu_1^{(n)}+a_2\nu_2^{(n)}+\ldots+a_m\nu_m^{(n)},$$

применим эмпирический закон устойчивости частот к каждой из $\nu_j^{(n)}$. Отсюда возникает приближенное равенство

$$n^{-1}(x_1+\ldots+x_n)\approx a_1p_1+a_2p_2+\ldots+a_mp_m=\mathbf{M}X,$$

выражающее возможность *статистической оценки* среднего значения случайной величины по результатам повторных независимых наблюдений. *Статистическая устойчивость средних арифметических* $n^{-1}(x_1+\ldots+x_n)$ не зависит от частных предположений о распределении сл. в. X, лишь бы существовало математическое ожидание, и представляет собой эмпирический закон, подтверждаемый всем опытом естествознания.

Математическое ожидание дискретной случайной величины, заданной на произвольном пространстве $(\Omega, \mathcal{A}, P)$, определяется с помощью формулы, которая использовалась выше для вычисления математического ожидания. Именно, пусть сл. в. X принимает значения $x_1, x_2, \ldots (x_i\neq x_j, i\neq j)$. Введем разбиение $\{B_j, j=1, 2, \ldots\}$ пространства Ω, полагая $B_j=\{\omega : X(\omega)=x_j\}$, и определим математическое ожидание сл. в. X как сумму ряда

$$\mathbf{M}X=\Sigma_{\{j\}}x_j\mathbf{P}(B_j)$$

в предположении, что он абсолютно сходится. Отметим, что если сл. в. X постоянна с вероятностью 1, то $\mathbf{M}X$ равно этой постоянной.

7.3. Задача. Сл. в. X принимает целые неотрицательные значения. Вывести формулу

$$\mathbf{M}X=\sum_{k=1}^{\infty}\mathbf{P}(\omega : X(\omega)\geqslant k).$$

$$\Big[\sum_{k=1}^{\infty}\mathbf{P}(\omega : X(\omega)\geqslant k)=\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\mathbf{P}(\omega : X(\omega)=n)=$$

$$=\sum_{n=1}^{\infty}\sum_{k=1}^{n}\mathbf{P}(\omega : X(\omega)=n)=\sum_{n=1}^{\infty}n\mathbf{P}(\omega : X(\omega)=n).$$

Отметим, что на первом шаге была использована счетная аддитивность вероятностной меры.]

7.4. Задача. Пусть $\{A_k, k=1, 2, \ldots\}$ — некоторое разбиение пространства элементарных событий Ω $(A_k\in\mathcal{A}, A_kA_l=\varnothing$ при

$k\neq l$, $\bigcup\limits_{\{k\}} A_k=\Omega$). Предположим, что сл. в. X постоянна на элементах разбиения: $X(\omega)=y_k$ при $\omega\in A_k$. Вывести формулу

$$\mathbf{M}X=\sum_{\{k\}} y_k \mathbf{P}(A_k).$$

[Пусть x_1, x_2, ... — набор различных значений дискретной сл. в. X, $B_j=\{\omega : X(\omega)=x_j\}$, $j=1, 2, \ldots$. Воспользовавшись представлениями

$$\mathbf{P}(B_j)=\sum_{\{k\}}\mathbf{P}(B_jA_k),\quad \mathbf{P}(A_k)=\sum_{\{j\}}\mathbf{P}(B_jA_k),$$

которые вытекают из счетной аддитивности меры (впрочем, второе равенство, ввиду $A_k\subseteq B_j$ при некотором j, тривиально), получаем

$$\mathbf{M}X=\sum_{\{j\}} x_j\mathbf{P}(B_j)=\sum_{\{j\}}\sum_{\{k\}} x_j\mathbf{P}(B_jA_k)=\sum_{\{j\}}\sum_{\{k\}} y_k\mathbf{P}(B_jA_k)=$$

$$=\sum_{\{k\}} y_k\sum_{\{j\}}\mathbf{P}(B_jA_k)=\sum_{\{k\}} y_k\mathbf{P}(A_k),$$

где мы учли, что при значениях j, k, для которых $\mathbf{P}(B_jA_k)>0$ (и, следовательно, $B_jA_k\neq\varnothing$), выполнено равенство $x_j=X(\omega)=y_k$ при $\omega\in B_jA_k$. Таким образом, если $\mathbf{M}X$ существует, то ряд $\Sigma y_k\mathbf{P}(A_k)$ абсолютно сходится и имеет место требуемое равенство. Проводя рассуждения в обратную сторону, получаем, что из абсолютной сходимости ряда $\Sigma y_k\mathbf{P}(A_k)$ вытекает существование $\mathbf{M}X$ и требуемое равенство.]

7.5. З а д а ч а. Вывести формулы $\mathbf{M}(cX)=c\mathbf{M}X$, $c=\mathrm{const}$, $\mathbf{M}(X+Y)=\mathbf{M}X+\mathbf{M}Y$, если $\mathbf{M}X$ и $\mathbf{M}Y$ существуют (линейность операции $\mathbf{M}X$).

[Пусть x_1, x_2, ..., y_1, y_2, ... — наборы значений сл. в. X, Y соответственно, $B_j=\{\omega : X(\omega)=x_j\}$, $A_k=\{\omega : Y(\omega)=y_k\}$, j, $k=1, 2, \ldots$. Сл. в. $X+Y$ постоянна на элементах разбиения $\{A_jB_k,\ j,\ k=1, 2, \ldots\}$ и принимает значение x_j+y_k при $\omega\in A_jB_k$. Применяя 7.4, получаем

$$\mathbf{M}(X+Y)=\sum_{\{j,k\}}(x_j+y_k)\mathbf{P}(A_jB_k)$$

в предположении, что ряд в правой части абсолютно сходится. Но поскольку $\mathbf{M}X$, $\mathbf{M}Y$ существуют, то из 7.4 вытекает, что

$$\mathbf{M}X=\sum_{\{j,k\}} x_j\mathbf{P}(A_jB_k),\quad \mathbf{M}Y=\sum_{\{j,k\}} y_j\mathbf{P}(A_jB_k),$$

причем ряды абсолютно сходятся. Складывая, получаем требуемый результат.]

7.6. Задача. Вычислить среднее значение биномиального и гипергеометрического распределений (см. 7.2), представив соответствующие случайные величины в виде суммы индикаторов.

[(I) Пусть $X_1, X_2, \ldots, X_n$ — последовательность испытаний Бернулли. Тогда сл. в. $S_n = X_1 + \ldots + X_n$ имеет биномиальное распределение, и так как $\mathbf{M}X_i = \mathbf{P}(\omega : X_i(\omega) = 1) = p$, то $\mathbf{M}S_n = np$. (II) Полагая $X_i = 1$, если i-й вытащенный шар в схеме повторного выбора без возвращения красный, имеем $\mathbf{M}X_i = \mathbf{P}(\omega : X_i(\omega) = 1) = p$, $\mathbf{M}S_n = np$, где $S_n = X_1 + \ldots + X_n$ — число красных шаров в выборке объема, n — имеет гипергеометрическое распределение, p — доля красных шаров.]

7.7. Задача. Вычислить математическое ожидание отрицательной биномиальной случайной величины, представив ее как сумму геометрически распределенных случайных величин (см. 4.6).

[Пусть $X_1, \ldots, X_r$ — независимые случайные величины с геометрическим распределением $\mathbf{P}(\omega : X_i(\omega) = k) = (1-p)^k p$, $k = 0$, 1 2, ... Тогда

$$\mathbf{M}X_i = \sum_{k=1}^{\infty} \mathbf{P}(\omega : X_i(\omega) \geqslant k) = \sum_{k=1}^{\infty} (1-p)^k = (1-p)/p,$$

откуда $\mathbf{M}S_r = r(1-p)/p$, где $S_r = X_1 + \ldots + X_r$ имеет отрицательное биномиальное распределение (см. 4.6).]

7.8*. Задача. Вывести формулу для вероятности объединения событий (*принцип включения — исключения*):

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{\{i_1 < \ldots < i_k\}} \mathbf{P}(A_{i_1} \ldots A_{i_k}),$$

вводя индикаторные сл. в. $I_{A_i}(\omega) = 1$ при $\omega \in A_i$ и пользуясь соотношениями $I_{AB} = I_A I_B$, $\mathbf{M}I_A = \mathbf{P}(A)$,

$$I_{\bar{A}} = 1 - I_A, \quad \overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \bar{A}_i.$$

$$\Big[\mathbf{P}\left(\bigcap_{i=1}^{n} \bar{A}_i\right) = \mathbf{M}\left(\prod_{i=1}^{n} I_{\bar{A}_i}\right) = \mathbf{M}\left(\prod_{i=1}^{n} (1 - I_{A_i})\right) =$$

$$= \mathbf{M}\left(1 - \sum_{k=1}^{n} (-1)^{k-1} \sum_{\{i_1 < \ldots < i_k\}} I_{A(i_1)} \ldots I_{A(i_k)} =\right.$$

$$= 1 - \mathbf{M} \sum_{k=1}^{n} (-1)^{k-1} \sum_{\{i_1 < \ldots < i_k\}} I_{A(i_1) \ldots A(i_k)}.$$

Воспользовавшись свойством линейности математического ожидания 7.5, получаем требуемый результат.]

7.9. Задача. Найти вероятность того, что при случайном размещении n писем по n конвертам хотя бы одно письмо попадет в свой конверт, воспользовавшись формулой 7.8.

[Пусть событие A_i означает, что письмо с номером i попало в нужный конверт. Легко подсчитать, что

$$\mathbf{P}(A_i)=n^{-1},\quad \mathbf{P}(A_iA_j)=(n(n-1))^{-1},\quad i<j,\quad \mathbf{P}(A_iA_jA_k)=(n(n-1)(n-2))^{-1}$$

и т. д., откуда

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{k=1}^{n}(-1)^{k-1}C_n^k(n(n-1)\ldots(n-k+1))^{-1}=\sum_{k=1}^{n}(-1)^{k-1}/k!.$$

Заметим, что при $n\to\infty$ эта вероятность стремится к $1-e^{-1}$, причем уже для небольших n близость к предельному значению весьма хорошая: при $n=5$ погрешность приближения составляет около 0,001, а при $n=7$ — около 0,00002.]

7.10. Задача. Сл. в. $Y=g(\mathbf{X})$ — функция набора дискретных сл. в. $\mathbf{X}=(X_1, \ldots, X_n)$. Вывести формулу

$$\mathbf{M}Y=\sum_{\{\mathbf{x}\}} g(\mathbf{x})\,p_{\mathbf{X}}(\mathbf{x}).$$

[Сл. в. Y принимает постоянное значение $g(\mathbf{x})$ на элементе $B_{\mathbf{x}}=\{\omega : \mathbf{X}(\omega)=\mathbf{x}\}$ разбиения, порождаемого сл. в. $\mathbf{X}$. Остается воспользоваться результатом 7.4.]

7.11. Задача. Сл. в. X, Y независимы. Показать, что $\mathbf{M}XY=\mathbf{M}X\cdot\mathbf{M}Y$, если $\mathbf{M}X$ и $\mathbf{M}Y$ существуют.

[Ввиду 7.10 имеем

$$\mathbf{M}XY=\sum_{\{x,y\}} xy\,p_{X,Y}(x,\ y)=\sum_{\{x,y\}} xy\,p_X(x)\,p_Y(y)$$

при условии, что ряд абсолютно сходится. Но ряды

$$\mathbf{M}X=\sum_{\{x\}} x\,p_X(x),\quad \mathbf{M}Y=\sum_{\{y\}} y\,p_Y(y)$$

абсолютно сходятся, откуда и получаем требуемый результат.]

7.12. Задача. Показать, что: (I) если $|X|\leqslant|Y|$ и существует $\mathbf{M}Y$, то существует $\mathbf{M}X$; (II) если $X\geqslant 0$, то $\mathbf{M}X\geqslant 0$; (III) если $X\leqslant Y$, то $\mathbf{M}X\leqslant\mathbf{M}Y$; (IV) $|\mathbf{M}X|\leqslant\mathbf{M}|X|$, (в (II)—(IV) предполагается, что $\mathbf{M}X$ и $\mathbf{M}Y$ существуют).

[Все свойства вытекают непосредственно из определения математического ожидания. Кроме того, (III) вытекает из (II) и 7.5, (IV) вытекает из (III) и неравенств $-|X|\leqslant X\leqslant|X|$. Отметим, что, как следует из определения, $\mathbf{M}|X|$ существует, если существует $\mathbf{M}X$.]

7.13*. Задача. Пусть

$$\mathcal{P}_X(s)=\sum_{k=0}^{\infty} s^k p_X(k)=\mathbf{M}s^x, \quad |s|\leqslant 1,$$

— производящая функция (п. ф.) сл. в. X. Показать, что $\mathbf{M}X= =\mathcal{P}_X'(1-)$.

[Дифференцируя при $|s|<1$, имеем

$$\mathcal{P}_X'(s)=\sum_{k=1}^{\infty} ks^{k-1}p_X(k).$$

Если $\mathbf{M}X$ существует, то ряд $\sum kp_X(k)$ сходится и существует $\mathcal{P}_X'(1-)=\lim\limits_{s\to 1-}\mathcal{P}_X'(s)=\mathbf{M}X$ (теорема Абеля). Обратно:

$$\lim_{s\to 1-}\sum_{k=1}^{\infty} ks^{k-1}p_X(k)\geqslant\sum_{k=1}^{N} kp_X(k),$$

следовательно, $\mathbf{M}X$ существует, далее используем прямое утверждение.]

7.14. Задача. Вывести соотношение $\mathcal{P}_{X+Y}(s)=\mathcal{P}_X(s)\mathcal{P}_Y(s)$ для п. ф. суммы $X+Y$ независимых сл. в. X, Y, пользуясь свойством математического ожидания произведения независимых случайных величин.

$$[\mathcal{P}_{X+Y}(s)=\mathbf{M}s^{X+Y}=\mathbf{M}s^X\cdot s^Y=\mathbf{M}s^X\mathbf{M}s^Y=\mathcal{P}_X(s)\mathcal{P}_Y(s).]$$

7.15. Задача. Сл. в. N с целыми неотрицательными значениями не зависит от последовательности сл. в. X_1, X_2, ... Обозначим S_N случайную величину, которая при каждом $\omega\in\Omega$, таком, что $N(\omega)>0$, принимает значение $X_1(\omega)+\ldots+X_{N(\omega)}(\omega)$, а при $N(\omega)=0$ — равна нулю. Предполагая, что $\mathbf{M}X_i=a$ при всех i, вывести формулу $\mathbf{M}S_N=\mathbf{M}N\cdot\mathbf{M}X_1$ (если $\mathbf{M}N$ существует).

$$\Big[\mathbf{M}S_N=\sum_{(x)} x\mathbf{P}(\omega:S_N(\omega)=x)=\sum_{(x)} x\sum_{n=0}^{\infty}\mathbf{P}(\omega:S_N(\omega)=x,\ N(\omega)=n)=$$

$$=\sum_{(x)} x\sum_{n=0}^{\infty}\mathbf{P}(\omega:S_n(\omega)=x,\ N(\omega)=n)=\sum_{n=0}^{\infty}\mathbf{P}(\omega:N(\omega)=n)\times$$

$$\times\sum_{(x)} x\mathbf{P}(\omega:S_n(\omega)=x)=\sum_{n=1}^{\infty} na\mathbf{P}(\omega:N(\omega)=n)=a\cdot\mathbf{M}N.\Big]$$

7.16. Задача. Сл. в. N, X_1, X_2, ... независимы, принимают целые неотрицательные значения, сл. в. X_i, $i=1, 2, \ldots$, одинаково распределены, сл. в. $S_N=X_1+\ldots+X_N$ при $N>0$ и $S_N=0$ при $N=0$. Показать, что п. ф. сл. в. S_N равна $\mathcal{P}_N(\mathcal{P}_{X_1}(t))$, где $\mathcal{P}_N(t)$, $\mathcal{P}_{X_1}(t)$ — п. ф. сл. в. N, X_1 соответственно.

[Аналогично 7.15 имеем

$$Mt^{S_N} = \sum_{k=0}^{\infty} t^k \mathbf{P}(\omega : S_N(\omega) = k) =$$

$$= \sum_{n=0}^{\infty} \mathbf{P}(\omega : N(\omega) = n) \sum_{k=0}^{\infty} t^k \mathbf{P}(\omega : S_n(\omega) = k).$$

Ввиду 7.14 внутренняя сумма равна

$$\mathscr{P}_{S_n}(t) = (\mathscr{P}_{X_1}(t))^n,$$

откуда и получаем требуемый результат.]

Перейдем к определению математического ожидания произвольной сл. в. $X(\omega)$, заданной на некотором вероятностном пространстве $(\Omega, \mathscr{A}, \mathbf{P})$. Напомним, что числовая функция $X(\omega)$ называется случайной величиной, если для любого промежутка $I_{a,b}$ множество $\{\omega : X(\omega) \in I_{a,b}\}$ является событием, т. е. принадлежит классу $\mathscr{A}$. Соотношение

$$\mathbf{P}_X(I_{a,b}) = \mathbf{P}(\omega : X(\omega) \in I_{a,b})$$

определяет вероятностную меру P_X на классе промежутков (в § 10 будет показано, что это соотношение сохраняется, если $I_{a,b}$ заменить на любое борелевское множество из R). Распределения вероятностей в R общего вида изучались в § 6, непрерывные распределения вероятностей на прямой были введены в § 3. При этом мы руководствовались соображением, что измерение величин, наблюдаемых при проведении любого, в том числе вероятностного, опыта производится с некоторой точностью и что любая непрерывная модель должна обладать «сколь угодно точным» дискретным приближением. Поделив числовую прямую точками kn^{-1}, $k=0, \pm 1, \pm 2, \ldots$, на отрезки длины n^{-1}, определим дискретное распределение вероятностей

$$p^{(n)}(kn^{-1}) = \mathbf{P}_X((kn^{-1}, (k+1)n^{-1}]), \quad k=0, \pm 1, \pm 2, \ldots,$$

стягивающее в точку kn^{-1} вероятностную массу, приписываемую распределением $\mathbf{P}_X$ промежутку $(kn^{-1}, (k+1)n^{-1}]$. Полагаем, что любая достойная интереса характеристика распределения вероятностей $\mathbf{P}_X$ должна приближаться соответствующей характеристикой дискретного распределения $p^{(n)}$. В частности, среднее значение распределения $p^{(n)}$

$$\sum_{k=-\infty}^{\infty} kn^{-1} p^{(n)}(kn^{-1}) = \sum_{k=-\infty}^{\infty} kn^{-1} \mathbf{P}(\omega : kn^{-1} < X(\omega) \leqslant (k+1)n^{-1}),$$

если оно существует, следует принять за приближение к математическому ожиданию сл. в. X и *определить* $\mathbf{M}X$ как предел математических ожиданий последовательности дискретных распределений $p^{(n)}$ при $n \to \infty$. Это же самое на языке случайных величин можно выразить следующим образом. Введем последовательность сл. в. $X^{(n)}(\omega)$, полагая

$$X^{(n)}(\omega) = kn^{-1} \Leftrightarrow kn^{-1} < X(\omega) \leqslant (k+1)n^{-1}, \quad k = 0, \pm 1, \pm 2, \ldots .$$

Математическое ожидание или *среднее значение* сл. в. X определим как предел

$$\lim_{n \to \infty} \mathbf{M}X^{(n)} = \mathbf{M}X$$

математических ожиданий

$$\mathbf{M}X^{(n)} = \sum_{k=-\infty}^{\infty} kn^{-1} p^{(n)}(kn^{-1})$$

дискретных сл. в. $X^{(n)}$.

7.17. Задача. Показать, что $|X^{(n_1)}(\omega) - X^{(n_2)}(\omega)| \leqslant \max \times \times (n_1^{-1}, n_2^{-1})$. Вывести отсюда, что: (I) если $\mathbf{M}X^{(n)}$ существует при каком-нибудь n, то существует при всех n; (II) последовательность $\mathbf{M}X^{(n)}$ фундаментальна и, следовательно, существует $\lim_{n \to \infty} \mathbf{M}X^{(n)}$, если только $\mathbf{M}X^{(n)}$ существует; (III) математическое ожидание ограниченной сл. в. существует; (IV) $|\mathbf{M}X^{(n)} - \mathbf{M}X| \leqslant n^{-1}$.

[Пусть для некоторого ω $X(\omega) \in (kn_1^{-1}, (k+1)n_1^{-1}]$ и $X(\omega) \in \in (ln_2^{-1}, (l+1)n_2^{-1}]$. В таком случае указанные промежутки имеют непустое пересечение, а следовательно, разность $kn_1^{-1} - ln_2^{-1} = = X^{(n_1)}(\omega) - X^{(n_2)}(\omega)$ между их левыми концами не превосходит по модулю длины наибольшего из них. (I) Из неравенства $|X^{(n)}(\omega)| \leqslant \leqslant |X^{(m)}(\omega)| + \max(n^{-1}, m^{-1})$ и задач 7.12, 7.5 заключаем, что $\mathbf{M}X^{(n)}$ существует, если существует $\mathbf{M}X^{(m)}$. (II) Применяя 7.12, имеем

$$|\mathbf{M}X^{(n_1)} - \mathbf{M}X^{(n_2)}| \leqslant \mathbf{M}|X^{(n_1)} - X^{(n_2)}| \leqslant \max(n_1^{-1}, n_2^{-1}),$$

так что последовательность $\mathbf{M}X^{(n)}$ фундаментальна. (IV) Переходя в полученном неравенстве к пределу при $n_2 \to \infty$, получаем требуемое неравенство.]

Специальный вид дискретных сл. в. $X^{(n)}$, равномерно приближающих сл. в. $X : |X^{(n)}(\omega) - X(\omega)| \leqslant n^{-1}$, которые участвуют в определении $\mathbf{M}X$, как показывает следующая задача, не является существенным.

7.18. Задача. Рассмотрим разбиение $T = \{t_i, i = 0, \pm 1, \pm 2, \ldots, t_i < t_{i+1}\}$ числовой прямой R и пусть $Q = \{s_i, i = 0, \pm 1, \pm 2, \ldots\}$ — множество точек $s_i \in [t_i, t_{i+1}]$, $i = 0, \pm 1, \ldots$. Диаметром разбиения T назовем число $d(T) = \sup_{(i)} (t_{i+1} - t_i)$. Образуем ряд

$$S_X(T, Q) = \sum_{i=-\infty}^{\infty} s_i \mathbf{P}_X((t_i, t_{i+1}]).$$

Показать, что: (I) если $d(T)<\infty$ и ряд $S_X(T, Q)$ абсолютно сходится, то абсолютно сходится ряд $S_X(T', Q')$ для любых T', Q', таких, что $d(T')<\infty$; (II) при $d(T)\to 0$ все абсолютно сходящиеся ряды $S_X(T, Q)$ сходятся к одному и тому же числу $\mathbf{M}X$.

[Любым $X(\omega)$, T, Q поставим в соответствие дискретную сл. в. $X_{T,Q}(\omega)$, полагая

$$X_{T,Q}(\omega)=s_i \Leftrightarrow t_i<X(\omega)\leqslant t_{i+1}, \quad i=0, \pm 1, \pm 2, \ldots .$$

Тогда для любых T_1, Q_1 и T_2, Q_2 выполнено неравенство

$$|X_{T_1,Q_1}(\omega)-X_{T_2,Q_2}(\omega)|\leqslant 2\max(d(T_1), d(T_2)),$$

откуда выводятся утверждения (I), (II), точно так же, как в 7.17.]

7.19. Задача. Допустим, что последовательность $X_n(\omega)$ дискретных случайных величин сходится равномерно (на Ω) к сл. в. $X(\omega)$. Предположим, что $\mathbf{M}X_n$ существует при всех $n=1, 2, \ldots$ и сходится к конечному пределу. Показать, что тогда $\mathbf{M}X=\lim\limits_{n\to\infty}\mathbf{M}X_n$. Вывести отсюда, что для дискретной сл. в. X определение $\mathbf{M}X$ как суммы ряда эквивалентно определению $\mathbf{M}X$ как предела $\mathbf{M}X^{(n)}$.

[Пусть последовательность сл. в. $X^{(n)}(\omega)=kn^{-1}$ при $kn^{-1}<X(\omega)\leqslant(k+1)n^{-1}$, $k=0, \pm 1, \ldots$, так что $|X^{(n)}(\omega)-X(\omega)|\leqslant n^{-1}$. По условию $\sup_{(\omega)}|X_n(\omega)-X(\omega)|\equiv\varepsilon_n\to 0$ при $n\to\infty$. В таком случае при всех ω $|X_n(\omega)-X^{(n)}(\omega)|\leqslant\varepsilon_n+n^{-1}$. Применяя 7.12, заключаем, что существует $\mathbf{M}X^{(n)}$ и

$$|\mathbf{M}X_n-\mathbf{M}X^{(n)}|\leqslant\varepsilon_n+n^{-1}.$$

Устремляя n к ∞, получаем, что последовательность $\mathbf{M}X^{(n)}$ сходится к $\mathbf{M}X$. Наконец, для дискретной сл. в. X положим $X_n=X$, $n=1, 2, \ldots$, и, применяя полученный результат, делаем вывод об эквивалентности определений.]

В задачах 7.20—7.25 устанавливается ряд элементарных свойств математического ожидания, которые для дискретных случайных величин получены в 7.5, 7.11, 7.12.

7.20. Задача. Показать, что операция математического ожидания линейна: $\mathbf{M}(aX+bY)=a\mathbf{M}X+b\mathbf{M}Y$, в частности $\mathbf{M}(aX+b)=a\mathbf{M}X+b$.

[Свойство $\mathbf{M}(cX)=c\mathbf{M}X$ очевидно. Если $\mathbf{M}X$ и $\mathbf{M}Y$ существуют, то $\mathbf{M}(X^{(n)}+Y^{(n)})=\mathbf{M}X^{(n)}+\mathbf{M}Y^{(n)}\to\mathbf{M}X+\mathbf{M}Y$ при $n\to\infty$, где $X^{(n)}$, $Y^{(n)}$ — последовательности дискретных случайных величин, участвующих в определении $\mathbf{M}X$, $\mathbf{M}Y$. Поскольку $X^{(n)}$, $Y^{(n)}$ равномерно сходятся к X, Y соответственно, то $X^{(n)}+Y^{(n)}$ равномерно сходятся к $X+Y$ и остается воспользоваться результатом 7.19.]

7.21. Задача. Положим $X_+(\omega)=X(\omega)$ при $X(\omega)\geqslant 0$ и $X_+(\omega)=0$ при $X(\omega)<0$; $X_-(\omega)=-X(\omega)$ при $X(\omega)\leqslant 0$ и $X_-(\omega)=0$ при $X(\omega)>0$. Показать, что если $\mathbf{M}X$ существует, то существуют $\mathbf{M}X_+$, $\mathbf{M}X_-$ и $\mathbf{M}X=\mathbf{M}X_+-\mathbf{M}X_-$, и обратно, если существуют $\mathbf{M}X_+$, $\mathbf{M}X_-$, то существует $\mathbf{M}X$.]

[Пусть $X^{(n)}(\omega)$ — последовательность дискретных сл. в., участвующих в определении $\mathbf{M}X$. Так как $0 \leqslant X_+^{(n)}$, $X_-^{(n)} \leqslant |X^{(n)}|$, то (см. 7.12) $\mathbf{M}X_+^{(n)}$, $\mathbf{M}X_-^{(n)}$ существуют, $\mathbf{M}X_+^{(n)} \to \mathbf{M}X_+$, $\mathbf{M}X_-^{(n)} \to \mathbf{M}X_-$ при $n \to \infty$, а из равенства $X_+^{(n)} - X_-^{(n)} = X^{(n)}$ получаем $\mathbf{M}X_+ - \mathbf{M}X_- = \mathbf{M}X$.]

7.22. Задача. Показать, что $\mathbf{M}X$ существует тогда и только тогда, когда существует $\mathbf{M}|X|$.

[Пусть $\mathbf{M}X$ существует. Применяя 7.21, заключаем, что существуют $\mathbf{M}X_+$, $\mathbf{M}X_-$, откуда $\mathbf{M}|X| = \mathbf{M}(X_+ + X_-) = \mathbf{M}X_+ + \mathbf{M}X_-$ также существует. Обратно, поскольку $|X^{(n)}| \to |X|$ равномерно, то, применяя 7.19, получаем $\mathbf{M}|X^{(n)}| \to \mathbf{M}|X|$ при $n \to \infty$.]

7.23. Задача. Показать, что: (I) если $|X| \leqslant |Y|$ и существует $\mathbf{M}Y$, то существует $\mathbf{M}X$; (II) если $X \geqslant 0$ и $\mathbf{M}X$ существует, то $\mathbf{M}X \geqslant 0$; (III) если $X \leqslant Y$ и $\mathbf{M}X$, $\mathbf{M}Y$ существуют, то $\mathbf{M}X \leqslant \mathbf{M}Y$; (IV) если $\mathbf{M}X$ существует, то $|\mathbf{M}X| \leqslant \mathbf{M}|X|$.

[(I) Из неравенства $|X| \leqslant |Y|$ вытекает $|X^{(n)}| \leqslant |Y^{(n)}|$ и, следовательно (см. 7.12), существует $\mathbf{M}X^{(n)}$. (II) При $X \geqslant 0$ имеем $X^{(n)} \geqslant 0$ и потому $\mathbf{M}X^{(n)} \geqslant 0$, откуда $\mathbf{M}X \geqslant 0$. (III) $0 \leqslant \mathbf{M}(Y - X) = \mathbf{M}Y - \mathbf{M}X$. (IV) $|\mathbf{M}X| = |\mathbf{M}X_+ - \mathbf{M}X_-| \leqslant \mathbf{M}X_+ + \mathbf{M}X_- = \mathbf{M}|X|$.]

7.24. Задача. Если X, Y независимы и ограничены (так что $\mathbf{M}X$, $\mathbf{M}Y$ существуют), то $\mathbf{M}XY = \mathbf{M}X \cdot \mathbf{M}Y$ (общий случай см. в 10.36).

[Пусть $X^{(n)}$, $Y^{(n)}$ — случайные величины, участвующие в определении $\mathbf{M}X$, $\mathbf{M}Y$. Легко видеть, что $X^{(n)}$ и $Y^{(n)}$ независимы, так что, применяя 7.11, получаем $\mathbf{M}X^{(n)}Y^{(n)} = \mathbf{M}X^{(n)}\mathbf{M}Y^{(n)}$. Переходя к пределу при $n \to \infty$ и учитывая 7.19, приходим к нужному результату.]

7.25. Задача. Положим $X_{a,b} = X \cdot I_{(\omega\,:\,a < X(\omega) \leqslant b)}$. Показать, что если $\mathbf{M}X$ существует, то $\mathbf{M}X_{a,b} \to \mathbf{M}X$ при $a \to -\infty$, $b \to \infty$ (разобрать случай $X \geqslant 0$, а затем использовать равенство $X = X_+ - X_-$).

[Если $X \geqslant 0$, то при $b_1 < b_2$ имеем $X_{0,b_1}(\omega) \leqslant X_{0,b_2}(\omega)$, $\mathbf{M}X_{0,b_1} \leqslant \mathbf{M}X_{0,b_2} \leqslant \mathbf{M}X$, так что существует $\lim\limits_{b \to \infty} \mathbf{M}X_{0,b} = \mu$. Покажем, что $\mu = \mathbf{M}X$. В неравенстве 7.17 (IV) для сл. в. $X_{0,b}$

$$|\mathbf{M}X_{0,b}^{(n)} - \mathbf{M}X_{0,b}| \leqslant n^{-1}, \quad \mathbf{M}X_{0,b}^{(n)} = \sum_{k=0}^{[nb]} kn^{-1}p^{(n)}(kn^{-1})$$

перейдем к пределу при $b \to \infty$ и получим

$$|\mathbf{M}X^{(n)} - \mu| \leqslant n^{-1}.$$

Устремляя n к ∞, приходим к требуемому равенству $\mu = \mathbf{M}X$. Для сл. в. $X \leqslant 0$ имеем тот же результат: $\mathbf{M}X_{a,0} \to \mathbf{M}X$ при $a \to -\infty$. В общем случае, взяв математическое ожидание от обеих частей равенства

$$X \cdot I_{(a < X \leqslant b)} = X_+ \cdot I_{(0 < X \leqslant b)} - X_- \cdot I_{(a < X \leqslant 0)}$$

и воспользовавшись аддитивностью, перейдем к пределу

$$\lim_{a\to-\infty,\, b\to\infty} \mathbf{M}X_{a,b}=\mathbf{M}X_+-\mathbf{M}X_-=\mathbf{M}X.]$$

7.26. Задача. Показать, что для сл. в. X с плотностью распределения $f(x)$ математическое ожидание существует, если $\int\limits_{-\infty}^{\infty} |x| f(x)\,dx<$ $<\infty$. В этом случае $\mathbf{M}X=\int\limits_{-\infty}^{\infty} xf(x)\,dx$.

[Пусть сл. в. X неотрицательна. Так как

$$\sum_{k=0}^{[nb]} kn^{-1}\int\limits_{k/n}^{(k+1)/n} f(x)\,dx\leqslant\sum_{k=0}^{[nb]}\int\limits_{k/n}^{(k+1)/n}(x+n^{-1})f(x)\,dx\leqslant\int\limits_{0}^{(b+1)/n} xf(x)\,dx,$$

то существует

$$\mathbf{M}X^{(n)}=\sum_{k=0}^{\infty} kn^{-1}\int\limits_{k/n}^{(k+1)/n} xf(x)\,dx$$

и, следовательно, существует $\mathbf{M}X$. Используя результат и обозначения задачи 7.25, имеем

$$\mathbf{M}X_{0,m}=\lim_{n\to\infty}\mathbf{M}X_{0,m}^{(n)}=\lim_{n\to\infty}\sum_{k=0}^{nm} kn^{-1}\int\limits_{k/n}^{(k+1)/n} f(x)\,dx=\int\limits_0^m xf(x)\,dx,$$

$$\mathbf{M}X=\lim_{m\to\infty}\mathbf{M}X_{0,m}=\int\limits_0^{\infty} xf(x)\,dx.$$

В общем случае используем разложение $X=X_+-X_-$ и применим полученный результат отдельно к X_+ и X_-, откуда

$$\mathbf{M}X=\mathbf{M}X_+-\mathbf{M}X_-=\lim_{b\to\infty}\int\limits_0^b xf(x)\,dx-\lim_{a\to-\infty}\int\limits_a^0 xf(x)\,dx.]$$

7.27. Задача. Сл. в. X имеет гамма-распределение с плотностью $f(x)=\Gamma(p)^{-1}x^{p-1}e^{-x}$, $x>0$ (см. 3.14 и далее). Вычислить $\mathbf{M}X$.

$$\Big[\int\limits_0^{\infty} x\Gamma(p)^{-1}x^{p-1}e^{-x}dx=\int\limits_0^{\infty}\Gamma(p+1)^{-1}x^pe^{-x}dx\cdot\Gamma(p+1)\Gamma(p)^{-1}=$$

$$=\Gamma(p+1)\Gamma(p)^{-1}=p\Gamma(p)\Gamma(p)^{-1}=p.$$

В частности, экспоненциальное распределение ($p=1$) имеет среднее 1.]

7.28. Задача. Сл. в. X имеет *бета-распределение* с плотностью

$$f(x)=\mathrm{B}(a,\,b)^{-1}x^{a-1}(1-x)^{b-1},\quad 0<x<1,\quad f(x)=0 \text{ при } x\overline{\in}(0,\,1),$$

где (см. 4.19)

$$\mathrm{B}(a,\ b)=\int_0^1 x^{a-1}(1-x)^{b-1}\,dx=\Gamma(a)\,\Gamma(b)/\Gamma(a+b),\quad a,\ b>0$$

— бета-функция Эйлера. Найти $\mathbf{M}X$.

$$\Big[\int_0^1 x\mathrm{B}(a,\ b)^{-1}x^{a-1}(1-x)^{b-1}\,dx=\int_0^1 \mathrm{B}(a+1,\ b)^{-1}x^{a}(1-x)^{b-1}\,dx\times$$

$$\times\,\mathrm{B}(a+1,\ b)\,\mathrm{B}(a,\ b)^{-1}=\mathrm{B}(a+1,\ b)\,\mathrm{B}(a,\ b)^{-1}=a/(a+b).$$

В частном случае равномерного на $[0,\ 1]$ распределения ($a=1$, $b=1$) имеем $\mathbf{M}X=1/2$.]

7.29. Задача. Сл. в. $X\geqslant 0$ имеет плотность $f(x)$. Показать, что $\mathbf{M}X^r=\int_0^\infty x^r f(x)\,dx$, $r>0$, если интеграл конечен.

$$\Big[\sum_{k=0}^{nm} kn^{-1}\,\mathbf{P}(\omega:kn^{-1}<X^r(\omega)\leqslant(k+1)\,n^{-1})=$$

$$=\sum_{k=0}^{nm}\ \int\limits_{\{x:\,kn^{-1}<x^r\leqslant(k+1)n^{-1}\}} kn^{-1}f(x)\,dx.$$

Заменяя под знаком интеграла kn^{-1} на x^r либо на x^r-n^{-1}, получим соответственно оценки сверху и снизу:

$$\int_0^{m+1/n} x^r f(x)\,dx,\qquad \int_0^{m+1/n} x^r f(x)\,dx-n^{-1}.$$

Устремляя m к ∞, получаем, что существует $\mathbf{M}Y^{(n)}$, $Y=X^r$, и

$$\int_0^\infty x^r f(x)\,dx-n^{-1}\leqslant \mathbf{M}Y^{(n)}\leqslant\int_0^\infty x^r f(x)\,dx.$$

Переходя к пределу при $n\to\infty$, получаем $\mathbf{M}X^r=\int_0^\infty x^r f(x)\,dx.\Big]$

7.30. Задача. Вычислить $\mathbf{M}X^r$, $r>0$, для гамма- и бета-распределений (см. 7.27, 7.28).

[Вычисления аналогичны проведенным в 7.27, 77.28, в частности

$$\int_0^\infty x^r\Gamma(p)^{-1}x^{p-1}e^{-x}dx=(p+r-1)(p+r-2)\ \ldots\ p.]$$

7.31. Задача. Пусть сл. в. X имеет стандартное нормальное распределение. Показать, что

$$MX^{2k-1}=0,\quad MX^{2k}=1\cdot 3\cdot 5\cdot\ \ldots\ \cdot(2k-1),\quad k=1,\ 2,\ \ldots .$$

$$[MX^r=\int_{-\infty}^{\infty} x^r (2\pi)^{-1/2}\exp(-x^2/2)\,dx.$$

При r нечетном подынтегральная функция нечетна, а при r четном — четна. Отсюда $MX^{2k-1}=0$ и

$$MX^{2n}=2\int_0^{\infty} x^{2k}(2\pi)^{-1/2}\exp(-x^2/2)\,dx=$$

$$=\pi^{-1/2}2^k\int_0^{\infty} y^{k-1/2}e^{-y}dy=\pi^{-1/2}\,2^k\,\Gamma(k+1/2),$$

где совершена замена переменного интегрирования $y=x^2/2$. Учитывая, что $\Gamma(p+1)=p\Gamma(p)$ и

$$\Gamma(1/2)=\int_0^{\infty} y^{-1/2}\,e^{-y}dy=\sqrt{2}\int_0^{\infty} e^{-x^2/2}\,dx=\sqrt{\pi},$$

получаем

$$MX^{2k}=2^k(k-1/2)(k-3/2)\,\ldots\,(1/2)=1\cdot 3\cdot 5\cdot\ \ldots\ \cdot(2k-1).]$$

7.32. Задача. Пусть сл. в. $X\geqslant 0$. Показать, что MX существует тогда и только тогда, когда сходится ряд $\sum\limits_{n=1}^{\infty} P(\omega : X(\omega)\geqslant n) = \mu$. При этом справедливы неравенства $\mu\leqslant MX\leqslant\mu+1$.

[Как следует из 7.17, MX существует тогда и только тогда, когда существует $MX^{(1)}$, где $X^{(1)}(\omega)=k$ при $k<X(\omega)\leqslant k+1$, $k=0,\ 1,\ 2,\ \ldots$. Применяя 7.3, получаем

$$MX^{(1)}=\sum_{k=1}^{\infty} P(\omega : X^{(1)}(\omega)\geqslant k)=\mu.$$

Остается заметить, что $X^{(1)}<X\leqslant X^{(1)}+1$ и, следовательно, $MX^{(1)}\leqslant MX\leqslant MX^{(1)}+1$.]

7.33. Задача. Вывести формулу

$$MX=\int_0^{\infty}(1-F(x))\,dx-\int_{-\infty}^{0} F(x)\,dx,\quad F(x)=P(\omega : X(\omega)\leqslant x),$$

причем MX существует тогда и только тогда, когда оба интеграла сходятся.

[Так как

$$F_+(x)\equiv \mathbf{P}(\omega : X_+(\omega)\leqslant x)=F(x),\quad x\geqslant 0,\quad F_+(x)=0,\quad x<0,$$

$$F_-(x-)=\mathbf{P}(\omega : X_-(\omega)<x)=1-F(-x),\quad x\geqslant 0,\quad F_-(x)=0,\quad x<0,$$

где $X=X_+-X_-$ (см. 7.21), то формула 7.33 для сл. в. X_+, X_- приобретает вид

$$\mathbf{M}X_+=\int\limits_0^\infty (1-F(x))\,dx,\quad \mathbf{M}X_-=\int\limits_0^\infty F(-x)\,dx=\int\limits_{-\infty}^0 F(x)\,dx.$$

Таким образом, достаточно рассмотреть сл. в. $X\geqslant 0$. Рассмотрим последовательность сл. в. $X^{(n)}$, участвующих в определении $\mathbf{M}X$, и покажем, что

$$\mathbf{M}X^{(n)}\to \int\limits_0^\infty (1-F(x))\,dx,\quad n\to\infty.$$

Сл. в. $nX^{(n)}$ принимает целые неотрицательные значения и потому (см. 7.3)

$$\mathbf{M}nX^{(n)}=n\mathbf{M}X^{(n)}=\sum_{k=1}^\infty \mathbf{P}(\omega : nX^{(n)}(\omega)\geqslant k).$$

Легко видеть, что

$$\sum_{k=2}^\infty (1-F(kn^{-1}))\,n^{-1}\leqslant \int\limits_{1/n}^\infty (1-F(x))\,dx\leqslant \sum_{k=1}^\infty (1-F(kn^{-1}))\,n^{-1}.$$

Отсюда вытекает, что существование $\mathbf{M}X^{(n)}$ эквивалентно конечности интеграла $\int\limits_0^\infty (1-F(x))\,dx$ и при этом последовательность

$$\mathbf{M}X^{(n)}=\sum_{k=1}^\infty \mathbf{P}(\omega : nX^{(n)}(\omega)>k-1)\,n^{-1}=\sum_{k=0}^\infty (1-F(kn^{-1}))\,n^{-1}$$

сходится к интегралу.]

7.34. Задача. Сл. в. X неотрицательна. Вывести формулу

$$\mathbf{M}X^\alpha=\int\limits_0^\infty \alpha x^{\alpha-1}(1-F(x))\,dx,\quad \alpha>0.$$

$$[F_\alpha(x)\equiv \mathbf{P}(\omega : X^\alpha(\omega)\leqslant x)=\mathbf{P}(\omega : X(\omega)\leqslant x^{1/\alpha})=F(x^{1/\alpha}),$$

так что, применяя 7.33, получаем

$$\mathbf{M}X^\alpha=\int\limits_0^\infty (1-F_\alpha(x))\,dx=\int\limits_0^\infty (1-F(x^{1/\alpha}))\,dx.$$

Замена переменного интегрирования $y=x^{1/\alpha}$ приводит к требуемому результату.]

Подобно тому как в анализе рассматривают бесконечные пределы, полезно обобщить понятие математического ожидания, допуская для $\mathsf{M}X$ значения $+\infty$ и $-\infty$. Именно, положим $\mathsf{M}X=+\infty$ для сл. в. $X\geqslant 0$, для которой ряд $\mathsf{M}X^{(n)}$ расходится (к $+\infty$), а также для произвольной сл. в. X, у которой $\mathsf{M}X_+=+\infty$, $\mathsf{M}X_-<+\infty$. Аналогично $\mathsf{M}X=-\infty$, если $\mathsf{M}X_+<\infty$, $\mathsf{M}X_-=+\infty$. Считаем $\mathsf{M}X$ определенным и в случае, когда оно принимает бесконечное значение, выделяя прежнее определение дополнительной оговоркой о конечности $\mathsf{M}X$.

§ 8.

ДИСПЕРСИЯ, КОВАРИАЦИЯ, СРЕДНЕКВАДРАТИЧЕСКОЕ РАССТОЯНИЕ

Операция $\mathsf{M}X$ усреднения значений сл. в. X, примененная к $Y=|X-a|^\alpha$, $\alpha\geqslant 1$, может служить характеристикой разброса значений сл. в. X относительно точки a. При условии, что $\mathsf{M}Y$ конечно, имеем (см. 7.34, где в качестве X надо взять $|X-a|$)

$$\mathbf{M}|X-a|^\alpha=\int_0^\infty \alpha x^{\alpha-1}\mathbf{P}(\omega:|X(\omega)-a|>x)\,dx\geqslant$$

$$\geqslant\int_0^t \alpha x^{\alpha-1}\mathbf{P}(\omega:|X(\omega)-a|>t)\,dx=t^\alpha\mathbf{P}(\omega:|X(\omega)-a|>t),$$

откуда вытекает важное неравенство

$$P(\omega:|X(\omega)-a|\geqslant t)\leqslant t^{-\alpha}\mathbf{M}|X-a|^\alpha,$$

оценивающее вероятностную массу (распределенную по закону $\mathbf{P}_X(B)=\mathbf{P}(\omega:X(\omega)\in B)$, $B\in\mathscr{B}$), лежащую за пределами $[t-a, t+a]$. Особенно интересен, как увидим, частный случай этого неравенства, когда $\alpha=2$ ($a=0$ или $a=\mathsf{M}X$). Его называют по имени русского ученого П. Л. Чебышева. Имея в виду аналогию между вероятностным распределением на прямой и распределением единичной массы, полезно отметить, что $\mathsf{M}X$ — координата центра масс с распределением $\mathbf{P}_X$, а $\mathsf{M}(X-a)^2$ — момент инерции линейно распределенной массы относительно точки a.

8.1. Задача. Показать, что если конечно $\mathsf{M}|X|^\alpha$, $\alpha\geqslant 1$, то конечно $\mathsf{M}|X-a|^\alpha$ при любом a.

[Для любых x, $y\geqslant 0$ имеем $(x+y)^\alpha\leqslant(2x)^\alpha$, если $x\geqslant y$ и $(x+y)^\alpha\leqslant(2y)^\alpha$, если $x\leqslant y$, так что $(x+y)^\alpha\leqslant 2^\alpha(x^\alpha+y^\alpha)$. Отсюда

$$\mathbf{M}|X-a|^\alpha\leqslant\mathbf{M}(|X|+|a|)^\alpha\leqslant 2^\alpha\,(\mathbf{M}|X|^\alpha+|a|^\alpha).]$$

8.2. Задача. Показать, что если конечно $\mathbf{M}|X|^{\alpha}$, $\alpha>0$, то конечно $\mathbf{M}|X|^{\beta}$, $0<\beta<\alpha$.

[Вытекает из неравенства $|X|^{\beta}\leqslant 1+|X|^{\alpha}$.]

8.3. Задача. Показать, что $\mathbf{M}(x-a)^2$ достигает минимума при $a=\mathbf{M}X$ (предполагается, что $\mathbf{M}X^2<\infty$).

[Из конечности $\mathbf{M}X^2$ вытекает конечность $\mathbf{M}X$, так что

$$\mathbf{M}(X-a)^2=\mathbf{M}(X^2-2Xa+a^2)=\mathbf{M}X^2-2a\mathbf{M}X+a^2.$$

Квадратичная функция обращается в минимум в точке $a=\mathbf{M}X$.]

Допустим, что $\mathbf{M}X^2<\infty$. Назовем *дисперсией* сл. в. X число

$$\mathbf{D}X=\mathbf{M}(X-\mathbf{M}X)^2.$$

Дисперсией распределения вероятностей $\mathbf{P}(B)$, $B\in\mathscr{B}$, на прямой назовем дисперсию случайной величины, распределение вероятностей которой есть $\mathbf{P}(B)$.

8.4. Задача. Вывести формулу $\mathbf{D}(X)=\mathbf{M}X^2-(\mathbf{M}X)^2$.

$$[\mathbf{M}(X-\mathbf{M}X)^2=\mathbf{M}(X^2-2X\mathbf{M}X+(\mathbf{M}X)^2)=\mathbf{M}X^2-(\mathbf{M}X)^2.]$$

8.5. Задача. Вычислить дисперсию биномиального распределения.

[Воспользовавшись тем, что $\mathbf{M}X^2=\mathbf{M}(X(X-1)+X)=\mathbf{M}X(X-1)+\mathbf{M}X$, запишем формулу для дисперсии в виде

$$\mathbf{D}X=\mathbf{M}X(X-1)+\mathbf{M}X-(\mathbf{M}X)^2.$$

В 7.1 показано, что

$$\mathbf{M}X=\sum_{k=0}^{n}kC_n^kp^k(1-p)^{n-k}=np.$$

Действуя аналогично, находим, что

$$\mathbf{M}X(X-1)=\sum_{k=0}^{n}k(k-1)C_n^kp^k(1-p)^{n-k}=n(n-1)p^2,$$

откуда $\mathbf{D}X=n(n-1)p^2+np-(np)^2=np(1-p)$.]

8.6*. Задача. Пусть $\mathscr{P}_X(s)$ — п. ф. сл. в. X. Показать, что

$$\mathbf{D}X=\mathscr{P}''_X(1-)+\mathscr{P}'_X(1-)-(\mathscr{P}'_X(1-))^2.$$

[При $|s|<1$ имеем (ср. 7.13):

$$\mathscr{P}''_X(s)=\sum_{k=2}^{\infty}k(k-1)s^{k-2}p_X(k)\to\mathbf{M}X(X-1),\ s\to 1-.$$

Остается воспользоваться формулой для дисперсии из 8.5 и тем, что $\mathscr{P}_X'(1-)=\mathbf{M}X$ (см. 7.13).]

8.7. Задача. Вычислить дисперсии отрицательного биномиального и пуассоновского распределений (п. ф. см. в 4.10).

[(I) $\mathcal{P}(s)=(1-p)^r(1-ps)^{-r}$, $\mathcal{P}'(s)=rp(1-p)^r(1-ps)^{-r-1}$, $\mathcal{P}''(s)=$ $=r(r+1)p^2(1-p)^r(1-ps)^{-r-2}$, откуда $\mathbf{D}X=r(r+1)p^2(1-p)^{-2}+$ $+rp(1-p)^{-1}-(rp(1-p)^{-1})^2=rp^2(1-p)^{-2}+rp(1-p)^{-1}=rp(1-p)^{-2}$.

(II) $\mathcal{P}(s)=\exp(-\lambda+\lambda s)$, $\mathcal{P}'(s)=\lambda\exp(-\lambda+\lambda s)$, $\mathcal{P}''(s)=\lambda^2\times$ $\times\exp(-\lambda+\lambda s)$, $\mathbf{D}X=\lambda^2+\lambda-\lambda^2=\lambda$. Подчеркнем, что $\mathbf{M}X=\mathbf{D}X$.]

8.8. Пример (закон больших чисел Бернулли). Обозначим S_n число успехов в n испытаниях Бернулли. Применяя неравенство Чебышева, получим

$$\mathbf{P}(\omega:|S_n(\omega)-\mathbf{M}S_n|\geqslant t)\leqslant t^{-2}\mathbf{D}S_n=t^{-2}np(1-p)$$

или

$$\mathbf{P}(\omega:|n^{-1}S_n(\omega)-p|\geqslant tn^{-1})\leqslant t^{-2}np(1-p).$$

Выбрав $t=\varepsilon n$, имеем

$$\mathbf{P}(\omega:|n^{-1}S_n(\omega)-p|\geqslant\varepsilon)\leqslant n^{-1}\varepsilon^{-2}p(1-p).$$

Таким образом, при любом $\varepsilon>0$ и $n\to\infty$

$$\mathbf{P}(\omega:|n^{-1}S_n(\omega)-p|\geqslant\varepsilon)\to 0.$$

Это соотношение называется *законом больших чисел* (ЗБЧ) Бернулли. ЗБЧ Бернулли представляет собой утверждение теории (относящееся к вероятностной модели повторных независимых испытаний), отражающее эмпирическую закономерность, выражаемую приближенным равенством $n^{-1}s_n\approx p$, где $n^{-1}s_n$ есть частота выпадения успеха.

8.9. Задача. Показать, что $\mathbf{P}(\omega:|r^{-1}Z_r(\omega)-p(1-p)^{-1}|\geqslant$ $\geqslant\varepsilon)\to 0$ при $r\to\infty$ и любом $\varepsilon>0$, где Z_r — число неудач, предшествующих r-му успеху в схеме Бернулли.

[По неравенству Чебышева

$$\mathbf{P}(\omega:|Z_r(\omega)-rp(1-p)^{-1}|\geqslant t)\leqslant t^{-2}rp(1-p)^{-2},$$

или

$$\mathbf{P}(\omega:|r^{-1}Z_r(\omega)-p(1-p)^{-1}|\geqslant\varepsilon)\leqslant r^{-1}\varepsilon^{-2}p(1-p)^{-2},$$

и остается устремить r к ∞.]

8.10. Задача. Сл. в. X_λ имеет распределение Пуассона с параметром λ. Показать, что при любом $\varepsilon>0$ и $\lambda\to\infty$

$$\mathbf{P}(\omega:|X_\lambda(\omega)-\lambda|\geqslant\lambda\varepsilon)\to 0.$$

[Вытекает из неравенства Чебышева. Отметим, что сл. в. $X_\lambda(\omega)$ определена, вообще говоря, на своем вероятностном пространстве, так что было бы правильнее писать $X_\lambda(\omega_\lambda)$, где ω_λ — точка пространства элементарных событий Ω_λ, на котором определена сл. в. X_λ. С этой точки зрения более естественно представление результата 8.10 в форме утверждения, касающегося семейства $\mathbf{P}_\lambda$ распределений вероятностей на прямой: $\mathbf{P}_\lambda(B)=\mathbf{P}(\omega:X(\omega)\in B)$. Именно

$$\mathbf{P}_\lambda(k:|k-\lambda|\geqslant\lambda\varepsilon)=\Sigma_{\{k:|k-\lambda|\geqslant\lambda\varepsilon\}}e^{-\lambda}\lambda^k/k!\to 0,\ \lambda\to\infty.]$$

8.11. Пример (теорема Вейерштрасса). Закон больших чисел Бернулли утверждает (см. 8.8), что сл. в. $n^{-1}S_n$, принимающая значения kn^{-1} с вероятностями $C_n^k p^k (1-p)^{n-k}$, $k=0, 1, \ldots, n$, имеет при больших n распределение, концентрирующееся вблизи точки p:

$$\mathbf{P}(\omega: |n^{-1}S_n(\omega)-p| < \delta) > 1 - n^{-1}\delta^{-2}p(1-p).$$

Возьмем произвольную непрерывную на отрезке $[0, 1]$ функцию $u(x)$ и образуем сл. в. $u(n^{-1}S_n(\omega)) \equiv U_n(\omega)$. Распределение вероятностей сл. в. U_n при больших n сосредоточено вблизи значения $u(p)$ и, кроме того, $\mathbf{M}U_n \to u(p)$ при $n \to \infty$. Действительно, пусть $A_{n,\delta} = \{\omega: |n^{-1}S_n(\omega)-p| < \delta\}$, $I(A)$ — индикатор события A, $c = \sup\{|u(x)|, 0 \leqslant x \leqslant 1\}$. По $\varepsilon > 0$ выберем $\delta > 0$ из условия равномерной непрерывности функции u. Тогда

$$|\mathbf{M}U_n - u(p)| = |\mathbf{M}(U_n - u(p))| \leqslant \mathbf{M}|U_n - u(p)| =$$

$$= \mathbf{M}(|U_n - u(p)| \cdot I(A_{n,\delta})) + \mathbf{M}(|U_n - u(p)| \cdot I(\bar{A}_{n,\delta})) \leqslant$$

$$\leqslant \mathbf{M}(\varepsilon I(A_{n,\delta})) + \mathbf{M}(2cI(\bar{A}_{n,\delta})) = \varepsilon \mathbf{M}I(A_{n,\delta}) + 2c\mathbf{M}I(\bar{A}_{n,\delta}) =$$

$$= \varepsilon \mathbf{P}(A_{n,\delta}) + 2c\mathbf{P}(\bar{A}_{n,\delta}) \leqslant \varepsilon + 2cn^{-1}\delta^{-2}p(1-p).$$

Выбирая $\varepsilon > 0$ сколь угодно малым и устремляя n к ∞, получаем требуемое. Положим

$$u_n(p) \equiv \mathbf{M}U_n = \sum_{k=0}^{n} u(kn^{-1}) C_n^k p^k (1-p)^{n-k}.$$

Проведенные оценки показывают, что $u_n(p) \to u(p)$ при $n \to \infty$ равномерно по $p \in [0, 1]$. Поскольку $u_n(p)$ — многочлен от p степени n, то полученное утверждение представляет собой теорему Вейерштрасса о равномерном приближении непрерывной на отрезке функции многочленами.

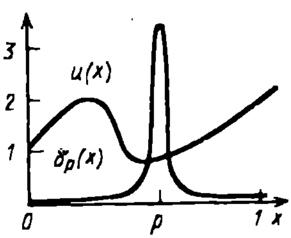

Рис. 27

Рис. 27 иллюстрирует доказанную теорему. Огибающая $\delta_p(x)$ функции дискретного аргумента $\delta_p(kn^{-1}) = nC_n^k p^k (1-p)^{n-k}$, $k=0, 1, \ldots, n$, является дельтаобразной с концентрацией вблизи точки p. При этом

$$\int_0^1 \delta_p(x)\,dx \approx \sum_{k=0}^{n} \delta_p(kn^{-1}) n^{-1} = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = 1,$$

$$u_n(p) = \sum_{k=0}^{n} u(kn^{-1}) \delta_p(kn^{-1}) n^{-1} \approx \int_0^1 \delta_p(x) u(x)\,dx \approx u(p).$$

8.12. Задача. Функция $u(x)$, $x \geqslant 0$, непрерывна и ограничена. Показать, что при каждом $\theta \geqslant 0$

$$u_n(\theta) = \sum_{k=0}^{\infty} u(kn^{-1}) e^{-n\theta} (n\theta)^k / k! \to u(\theta), \quad n \to \infty \quad (u_n(0) = u(0)),$$

причем сходимость равномерна на любом конечном промежутке. [Введем пуассоновскую сл. в. X_λ с параметром λ. Тогда $u_n(\theta) = \mathbf{M}u(n^{-1}X_{n\theta})$ и, действуя как в 8.11, находим

$$|u_n(\theta) - u(\theta)| \leqslant \varepsilon \mathbf{P}(\omega : |n^{-1}X_{n\theta}(\omega) - \theta| < \delta) +$$
$$+ 2c\mathbf{P}(\omega : |n^{-1}X_{n\theta} - \theta| \geqslant \delta), \; c = \sup\{|u(x)|, \; x \geqslant 0\}.$$

По неравенству Чебышева

$$\mathbf{P}(\omega : |X_{n\theta}(\omega) - n\theta| \geqslant \delta n) \leqslant n\theta/(\delta n)^2 = \theta/(\delta^2 n),$$

так что

$$|u_n(\theta) - u(\theta)| \leqslant \varepsilon + 2c\theta\delta^{-2}n^{-1},$$

откуда и следует требуемый результат.]

8.13. Задача. Вычислить дисперсии стандартного нормального и гамма-распределений.
[Вычисления проведем по формуле $\mathbf{D}X = \mathbf{M}X^2 - (\mathbf{M}X)^2$. Применяя 7.30 и 7.31, получаем $\mathbf{D}X = \mathbf{M}X^2 = 1$, $\mathbf{D}X = (p+1)p - p^2 = p$.]

8.14. Задача. Показать, что $\mathbf{D}(aX+b) = a^2\mathbf{D}X$, где a, b — постоянные.

[$\mathbf{D}(aX+b) = \mathbf{M}((aX+b) - \mathbf{M}(aX+b))^2 = \mathbf{M}(a^2(X - \mathbf{M}X)^2) = a^2\mathbf{D}X$.]

8.15. Задача. Сл. в. X имеет нормальное распределение с параметром сдвига a и параметром масштаба $\sigma > 0$. Показать, что $\mathbf{M}X = a$, $\mathbf{D}X = \sigma^2$.
[Плотность распределения сл. в. X равна $(2\pi)^{-1/2}\sigma^{-1}\varphi(\sigma^{-1}(x-a))$, $\varphi(x) = \exp(-x^2/2)$. Сл. в. $Y = (X-a)/\sigma$ имеет плотность распределения $\varphi(x)$, так что $0 = \mathbf{M}Y = \sigma^{-1}\mathbf{M}(X-a) = \sigma^{-1}(\mathbf{M}X - a)$, т. е. $\mathbf{M}X = a$, и $1 = \mathbf{D}Y = \sigma^{-2}\mathbf{D}X$, т. е. $\mathbf{D}X = \sigma^2$.]

8.16. Задача. Показать, что $\mathbf{D}X = 0$ тогда и только тогда, когда $X(\omega) = \mathrm{const}$ с вероятностью 1.
[В одну сторону утверждение очевидно. Допустим, что $\mathbf{D}X = 0$, и запишем неравенство Чебышева

$$\mathbf{P}(\omega : |X(\omega) - \mathbf{M}X| \geqslant t) \leqslant t^{-2}\mathbf{D}X = 0, \qquad t > 0.$$

Требуемое утверждение вытекает из свойства непрерывности вероятностной меры $\mathbf{P}$ (см. § 6):

$$\mathbf{P}(\omega : |X(\omega) - \mathbf{M}X| > 0) = \lim_{n \to \infty} \mathbf{P}(\omega : |X(\omega) - \mathbf{M}X| \geqslant n^{-1}) = 0.$$

Другое доказательство получается из следующего рассуждения. Пусть сл. в. X неотрицательна, $X^{(n)}$ — дискретная сл. в., участвующая в определении $\mathbf{M}X$: $X^{(n)}(\omega) = kn^{-1}$ при $kn^{-1} < X(\omega) \leqslant (k+1)n^{-1}$, $k = 0, \pm 1, \pm 2, \ldots$. Предположим, что $\mathbf{M}X = 0$. Из неравенства $0 \leqslant X^{(n)}(\omega) \leqslant X(\omega)$ заключаем, что $\mathbf{M}X^{(n)} = 0$. Отсюда, очевидно, вытекает, что $X^{(n)}(\omega)=0$ с вероятностью 1. Из неравенства $|X^{(n)}(\omega)-X(\omega)|<n^{-1}$ выводим, что и $X(\omega)=0$ с вероятностью 1.]

Если сл. в. $X(\omega)$, $Y(\omega)$ независимы, то (см. 7.24, 10.36)

$$\mathbf{M}XY=\mathbf{M}X\cdot\mathbf{M}Y.$$

Нетрудно понять, что обратное утверждение не справедливо, т. е. из равенства $\mathbf{M}XY=\mathbf{M}X\cdot\mathbf{M}Y$, вообще говоря, не следует независимость сл. в. X, Y. Так, для дискретных сл. в. X, Y из равенства

$$\sum_{(x,y)} xyp_{X,Y}(x, y)=\sum_{(x)} xp_X(x)\sum_{(y)} yp_Y(y)$$

нельзя в общем случае вывести, что $p_{X,Y}(x, y)=p_X(x)p_Y(y)$. Для индикаторных сл. в. I_A, I_B имеем

$$\mathbf{M}I_A=\mathbf{P}(A),\mathbf{M}I_B=\mathbf{P}(B),\quad \mathbf{M}I_AI_B=\mathbf{M}I_{AB}=\mathbf{P}(AB),$$

и поэтому равенство $\mathbf{M}I_AI_B=\mathbf{M}I_A\cdot\mathbf{M}I_B$ приобретет вид

$$\mathbf{P}(AB)=\mathbf{P}(A)\cdot\mathbf{P}(B),$$

значит, сл. в. I_A, I_B независимы. В общем случае разность $\mathbf{M}XY-\mathbf{M}X\cdot\mathbf{M}Y$ может быть использована в качестве некоторой характеристики зависимости между сл. в. X и Y, она называется *ковариацией* сл. в. X, Y и обозначается $\operatorname{cov}(X, Y)$. Отметим, что $\operatorname{cov}(X, X)=\mathbf{D}X$.

8.17. Задача. Если $\mathbf{M}X^2$ и $\mathbf{M}Y^2$ конечны, то конечно $\mathbf{M}XY$ и, следовательно, определена $\operatorname{cov}(X, Y)$.

[Следует из неравенства $2|X|\cdot|Y|\leqslant|X|^2+|Y|^2$.]

8.18. Задача. Вывести формулу $\operatorname{cov}(X, Y)=\mathbf{M}((X-\mathbf{M}X)(Y-\mathbf{M}Y))$.

[$\mathbf{M}((X-\mathbf{M}X)(Y-\mathbf{M}Y))=\mathbf{M}(XY-Y\mathbf{M}X-X\mathbf{M}Y+\mathbf{M}X\,\mathbf{M}Y)=\mathbf{M}XY-\mathbf{M}Y\cdot\mathbf{M}X-\mathbf{M}X\cdot\mathbf{M}Y+\mathbf{M}X\cdot\mathbf{M}Y=\mathbf{M}XY-\mathbf{M}X\cdot\mathbf{M}Y$.]

8.19. Задача. Показать, что для любых постоянных a, b, c, d справедливо равенство $\operatorname{cov}(aX+b, cY+d)=ac\operatorname{cov}(X, Y)$.

[$\operatorname{cov}(aX+b, cY+d)=\mathbf{M}(((aX+b)-\mathbf{M}(aX+b))((cY+d)-\mathbf{M}(cY+d)))=\mathbf{M}(a(X-\mathbf{M}X)\cdot c(Y-\mathbf{M}Y))=ac\mathbf{M}((X-\mathbf{M}X)\cdot(Y-\mathbf{M}Y))$.]

Как видно из 8.19, ковариация случайных величин не зависит от выбора начала отсчета, но зависит от выбора масштаба при измерениях, приводящих к рассматриваемым величинам. Чтобы получить характеристику зависимости, инвариантную относительно масштабных преобразований, поделим ковариацию X, Y на $(\mathbf{D}X\cdot\mathbf{D}Y)^{1/2}$. Полученная характеристика называется *коэффициентом корреляции*:

$$\rho(X, Y)=\operatorname{cov}(X, Y)/\sqrt{\mathbf{D}X\cdot\mathbf{D}Y}.$$

Говорят, что сл. в. $X^*=(X-\mathbf{M}X)/\sqrt{\mathbf{D}X}$ получена из сл. в. X центрированием (математическим ожиданием) и нормированием (корнем квадратным из дисперсии). Из 8.19 вытекает, что

$$\rho(X, Y)=\operatorname{cov}(X^*, Y^*).$$

8.20. Задача. Вывести неравенство Коши—Буняковского—Шварца

$$(\mathbf{M}XY)^2 \leqslant \mathbf{M}X^2 \cdot \mathbf{M}Y^2,$$

причем равенство достигается тогда и только тогда, когда с вероятностью 1 $aX(\omega)+bY(\omega)=0$ для некоторых постоянных a, b, не равных нулю одновременно.

[Пусть a, b — некоторые постоянные. Тогда

$$0 \leqslant \mathbf{M}(aX+bY)^2 = a^2\mathbf{M}X^2 + 2ab\mathbf{M}(XY) + b^2\mathbf{M}Y^2.$$

Рассматривая выражение справа при $\mathbf{M}X^2>0$ как квадратный трехчлен по a, выводим, что его дискриминант неположителен:

$$4(\mathbf{M}XY)^2 - 4\mathbf{M}X^2 \cdot \mathbf{M}Y^2 \leqslant 0.$$

К такому же заключению мы приходим, если $\mathbf{M}Y^2>0$. В случае, когда $\mathbf{M}X^2=\mathbf{M}Y^2=0$, сл. в. $X(\omega)=0$, $Y(\omega)=0$ с вероятностью 1, и неравенство тривиально. Равенство (см. 8.16) $\mathbf{P}(\omega: aX(\omega)+bY(\omega)=0)=1$ при $a^2+b^2>0$, $\mathbf{M}X^2+\mathbf{M}Y^2>0$ означает, что квадратный трехчлен имеет корень. Это эквивалентно равенству нулю дискриминанта.]

8.21. Задача. Вывести из 8.20, что $|\rho(X, Y)| \leqslant 1$, причем $|\rho(X, Y)|=1$ тогда и только тогда, когда с вероятностью 1 $aX(\omega)+bY(\omega)+c=0$ для некоторых постоянных a, b, c, $a^2+b^2>0$. (Предполагается, что $\mathbf{D}X$, $\mathbf{D}Y>0$, т. е. коэффициент корреляции определен.)

[Подставляя в 8.20 X^*, Y^* вместо X, Y, получаем

$$\rho(X^*, Y^*) \leqslant 1,\quad X^*=(X-\mathbf{M}X)/\sqrt{\mathbf{D}X},\quad Y^*=(Y-\mathbf{M}Y)/\sqrt{\mathbf{D}Y},$$

причем равенство имеет место, когда при некоторых a, b, $a^2+b^2>0$ с вероятностью 1 $aX^*(\omega)+bY^*(\omega)=0$, т. е. $a(X-\mathbf{M}X)+b(Y-\mathbf{M}Y)=0$.]

Одно из направлений в теории вероятностей изучает случайные величины на вероятностном пространстве $(\Omega, \mathscr{B}, \mathbf{P})$ как элементы линейного пространства функций на Ω с нормой

$$\|X\| = (\mathbf{M}X^2)^{1/2},$$

определенной для множества H всех случайных величин $X(\omega)$ с конечным математическим ожиданием квадрата. При этом отождествляем любые сл. в. $X(\omega)$, $Y(\omega)$, для которых $\|X-Y\|=0$, т. е. $\mathbf{P}(\omega: X(\omega)\neq Y(\omega))=0$. В H определяют скалярное произведение

$$\langle X, Y\rangle = \mathbf{M}XY,\quad \|X\|^2 = \langle X, X\rangle,$$

которое, очевидно, удовлетворяет необходимым свойствам:

1. $\langle X, X\rangle \geqslant 0$, причем $\|X\|=0 \Leftrightarrow \mathbf{P}(\omega: X(\omega)=0)=1$;
2. $\langle X, Y\rangle = \langle Y, X\rangle$;
3. $\langle a_1X_1+a_2X_2, Y\rangle = a_1\langle X_1, Y\rangle + a_2\langle X_2, Y\rangle$.

Расстояние $\|X-Y\|$ между элементами X, $Y \in H$ называется *среднеквадратическим расстоянием между* сл. в. $X(\omega)$, $Y(\omega)$.

Коэффициент корреляции

$$\rho(X, Y) = \langle X - \mathrm{M}X, Y - \mathrm{M}Y \rangle / (\|X - \mathrm{M}X\| \cdot \|Y - \mathrm{M}Y\|)$$

между сл. в. X, Y имеет геометрический смысл косинуса угла между векторами $X-\mathrm{M}X$ и $Y-\mathrm{M}Y$. Если $\rho(X, Y)=0$, то векторы $X-\mathrm{M}X$, $Y-\mathrm{M}Y$ *ортогональны*, а сл. в. $X(\omega)$, $Y(\omega)$ называют при этом *некоррелированными*. Независимые случайные величины некоррелированы: $\mathrm{M}XY=\mathrm{M}X\cdot\mathrm{M}Y$, обратное в общем случае неверно. Например, для сл. в. X, распределенной по стандартному нормальному закону, имеем $\mathrm{M}X=\mathrm{M}X^3=0$, так что $\mathrm{M}(X\cdot X^2)=\mathrm{M}X^3=0=\mathrm{M}X\cdot\mathrm{M}X^2$ и X, X^2 некоррелированы, хотя и связаны функциональной зависимостью. При $|\rho(X, Y)|=1$ сл. в. X, Y связаны линейной зависимостью, а в промежуточном случае $0<|\rho(X, Y)|<1$ между X и Y имеется связь, при этом коэффициент $\rho(X, Y)$ характеризует меру линейной связи между X и Y.

8.22. Задача. Найти коэффициенты a, b из условия минимума среднеквадратического расстояния $\|Y-(aX+b)\|$.

[Ввиду 8.3 минимум по b выражения $\mathrm{M}(Y-aX-b)^2$ достигается при $b=\mathrm{M}(Y-aX)=\mathrm{M}Y-a\mathrm{M}X$. Полагая $Y_0=Y-\mathrm{M}Y$, $X_0=X-\mathrm{M}X$, перепишем рассматриваемое выражение в виде

$$\mathrm{M}(Y_0-aX_0)^2=\mathrm{M}Y_0^2-2a\mathrm{M}X_0Y_0+a^2\mathrm{M}X_0^2.$$

Условие минимума по a дает

$$\widehat{a}=\mathrm{M}X_0Y_0/\mathrm{M}X_0^2=\mathrm{cov}(X, Y)/\mathrm{D}X, \quad \widehat{b}=\mathrm{M}Y-\widehat{a}\mathrm{M}X.]$$

Сл. в. $\widehat{Y}=\widehat{a}X+\widehat{b}$ — ближайшая к сл. в. Y линейная функция от сл. в. X. Говорят, что формула $\widehat{Y}=a\widehat{X}+\widehat{b}$ дает *наилучшую в среднеквадратичном линейную оценку* сл. в. Y. Если сл. в. X может быть измерена при проведении опыта, а сл. в. Y не допускает непосредственного наблюдения, но зато известны вероятностные характеристики $\mathrm{M}X$, $\mathrm{M}X^2$, $\mathrm{M}Y$, $\mathrm{M}Y^2$, $\mathrm{M}XY$, то $\widehat{Y}$ служит оценкой для Y, вычисленной по результату наблюдений за сл. в. X.

При геометрической интерпретации вектор $\widehat{Y}=\widehat{a}X+\widehat{b}$, найденный в 8.22, представляет собой проекцию вектора Y на линейное подпространство в H, натянутое на векторы X, 1. Коэффициенты $\widehat{a}$, $\widehat{b}$ можно было бы поэтому найти из соотношений ортогональности

$$\langle Y-aX-b, X\rangle=0, \quad \langle Y-aX-b, 1\rangle=0,$$

или

$$\langle Y, X\rangle-a\langle X, X\rangle-b\langle 1, X\rangle=0, \quad \langle Y, 1\rangle-a\langle X, 1\rangle-\langle b, 1\rangle=0,$$

т. е.

$$\mathrm{M}XY-a\mathrm{M}X^2-b\mathrm{M}X=0, \quad \mathrm{M}Y-a\mathrm{M}X-b=0,$$

откуда

$$\widehat{b}=\mathrm{M}Y-\widehat{a}\mathrm{M}X, \quad \widehat{a}=(\mathrm{M}XY-\mathrm{M}X\cdot\mathrm{M}Y)/(\mathrm{M}X^2-(\mathrm{M}X)^2).$$

Записав $Y=\widehat{Y}+(Y-\widehat{Y})$, имеем представление сл. в. Y в виде суммы сл. в. $\widehat{Y}$, линейно выражающейся через X, и сл. в. $Y-\widehat{Y}$, некоррелированной с X. В частности, для центрированных и нормированных сл. в. X^*, Y^* имеем

$$\widehat{Y}^*=\rho(X,\ Y)\cdot X^* \text{ и } Y^*=\rho(X,\ Y)\cdot X^*+(Y^*-\rho(X,\ Y)\cdot X^*).$$

8.23. Задача. Вывести следующую формулу дисперсии суммы случайных величин:

$$\mathbf{D}(X_1+\ldots+X_n)=\sum_{i=1}^{n}\mathbf{D}X_i+2\sum_{\{1\leqslant i<j\leqslant n\}}\operatorname{cov}(X_i,\ X_j).$$

В частности, если сл. в. $X_1, \ldots, X_n$ попарно некоррелированы, то дисперсия суммы равна сумме дисперсий.

[Ввиду 8.14, 8.19 можно положить $\mathbf{M}X_i=0$, $i=1, \ldots, n$ (заменив X_i на $X_i-\mathbf{M}X_i$), так что

$$\mathbf{D}(X_1+\ldots+X_n)=\mathbf{M}(X_1+\ldots+X_n)^2=\mathbf{M}\sum_{\{i,j\}}X_i\,X_j=$$

$$=\sum_{\{i\}}\mathbf{M}X_i^2+2\sum_{\{i<j\}}\mathbf{M}X_iX_j.]$$

8.24. Задача. Вычислить дисперсии биномиального и гипергеометрического распределений, используя представления соответствующих случайных величин в виде суммы.

[(I) Биномиальная сл. в. $S_n=X_1+\ldots+X_n$, где X_i, $i=1, \ldots, n$, — независимые индикаторы, $\mathbf{P}(\omega: X_i(\omega)=1)=p$, $i=1, \ldots, n$. Имеем

$$\mathbf{M}X_1=p,\ \mathbf{M}X_1^2=p,\ \mathbf{D}X_1=p-p^2,\ \mathbf{D}S_n=np(1-p).$$

(II) Гипергеометрическая сл. в. S_n — число красных шаров в выборке без возвращения из урны объема N с M красными шарами — представима в виде суммы $X_1+\ldots+X_n$, где X_i — индикатор события, что i-й вытащенный шар красный. Легко видеть, что

$$\mathbf{P}(\omega: X_i(\omega)=1)=M/N,\quad \mathbf{P}(\omega: X_i(\omega)=X_j(\omega)=1)=$$
$$=M(M-1)/(N(N-1)),$$

откуда

$$\mathbf{M}X_i=M/N,\ \mathbf{D}X_i=(M/N)(1-M/N),\ \operatorname{cov}(X_i,\ X_j)=M(M-1)/(N(N-1))-$$
$$-(M/N)^2,\ \mathbf{D}S_n=n(M/N)(1-M/N)+n(n-1)(M(M-1)/(N(N-1))-$$
$$-(M/N)^2)=n(M/N)(1-M/N)(1-(n-1)/(N-1)).]$$

8.25. Задача. Используя неравенство Чебышева и 8.23, вывести закон больших чисел в форме: пусть сл. в. $X_1, X_2, \ldots$ независимы и имеют одинаковые средние и дисперсии, тогда при $n \to \infty$ и любом $\varepsilon > 0$

$$\mathbf{P}(\omega : |n^{-1}(X_1(\omega) + \ldots + X_n(\omega)) - \mathbf{M}X_1| > \varepsilon) \to 0.$$

[Так как $\mathbf{D}(X_1 + \ldots + X_n) = n\mathbf{D}X_1$, то оценка сверху для рассматриваемой вероятности в неравенстве Чебышева равна $\mathbf{D}X_1/(n\varepsilon^2) \to 0$ при $n \to \infty$.]

В экспериментальных науках среднее арифметическое $\bar{x} = n^{-1}(x_1 + \ldots + x_n)$ из результатов $x_1, \ldots, x_n$ повторных измерений некоторой величины a рассматривают как более точное приближение к истинному значению a этой величины по сравнению с отдельным измерением. Вероятностная модель измерений дается последовательностью сл. в. $X_1, \ldots, X_n$. Если измерения не содержат систематической ошибки, т. е. отклонение $x_i - a$ объясняется чисто случайными погрешностями, то следует предположить, что $\mathbf{M}(X_i - a) = 0$, т. е. $\mathbf{M}X_i = a$. Если к тому же измерения независимы и одинаково точны: $\mathbf{D}X_i = \sigma^2$, $i = 1, \ldots, n$, то ЗБЧ 8.25 позволяет объяснить экспериментально наблюдаемую закономерность о стабилизации с ростом n средних арифметических $\bar{x} = n^{-1}(x_1 + \ldots + x_n)$ вблизи a. Качество приближения к a оценки $S(x_1, \ldots, x_n) = \bar{x}$ при конечных n можно характеризовать неравенством Чебышева

$$\mathbf{P}(\omega : |S(X_1(\omega), \ldots, X_n(\omega)) - a| > \varepsilon) \leqslant \varepsilon^{-2}\mathbf{D}S = \varepsilon^{-2}\sigma^2 n^{-1},$$

оценивающим степень концентрации распределения вероятностей сл. в. $S(X_1, \ldots, X_n)$ вокруг точки a. Эта оценка прямо зависит от дисперсии $\mathbf{D}S$, и, таким образом, точность приближения $S(x_1, \ldots, x_n) \approx a$ можно оценивать одним числом $\mathbf{D}S$. Как показывает следующая задача, не всегда среднее арифметическое из наблюдений — наилучший способ использования данных для оценивания математического ожидания.

8.26. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и равномерно распределены на отрезке $[0, \theta]$, $X_{(n)} = \max(X_1, \ldots, X_n)$. Найти $\mathbf{M}X_{(n)}$, $\mathbf{D}X_{(n)}$.

[Ф. р. $F_n(x)$ сл. в. $X_{(n)}$ и ее плотность $f_n(x)$ равны

$$F_n(x) = (x/\theta)^n,\ f_n(x) = \theta^{-n}nx^{n-1},\quad 0 \leqslant x \leqslant \theta,$$

откуда

$$\mathbf{M}X_{(n)} = \int_0^\theta x\theta^{-n}nx^{n-1}\,dx = \theta(1 - 1/n),\ \mathbf{M}X_{(n)}^2 = \theta^2 n/(n+2),$$

$$\mathbf{D}X_{(n)} = \mathbf{M}X_{(n)}^2 - (\mathbf{M}X_{(n)})^2 = \theta^2 n/((n+2)(n+1)^2).$$

Отсюда следует, что сл. в. $Y_n = 2^{-1}(1 + 1/n)^{-1}X_{(n)}$ имеет среднее

$$\mathbf{M}Y_n = \theta/2 = \mathbf{M}X_1,$$

т. е. оценка $2^{-1}(1 + 1/n)^{-1}\max(x_1, \ldots, x_n)$ величины $\theta/2$ — математического ожидания отдельного измерения, лишена систематической ошибки. Ее дисперсия равна

$$\mathbf{D}Y_n = 2^{-2}(1+1/n)^{-2}\mathbf{D}X_{(n)} = \theta^2/(4n(n+2)).$$

Среднее арифметическое $n^{-1}(X_1+\ldots+X_n)$ имеет дисперсию

$$n^{-1}\mathbf{D}X_1 = n^{-1}\left(\int\limits_0^\theta x^2\theta^{-1}\,dx - \left(\int\limits_0^\theta x\theta^{-1}\,dx\right)^2\right) = \theta^2/(12n).$$

Отношение

$$\mathbf{D}Y_n/\mathbf{D}(n^{-1}(X_1+\ldots+X_n)) = 3/(n+2)$$

меньше 1 при $n>1$ и стремится к 0 при $n\to\infty$.]

8.27. Задача. Допустим, что времена до отказа n идентичных изделий, поставленных на испытания, представляются n независимыми сл. в. $X_1, \ldots, X_n$, имеющими показательную плотность $\lambda e^{-\lambda x}$, $x\geqslant 0$. Предположим, что испытания продолжаются до r-го отказа ($r\leqslant n$), так что результаты наблюдений можно описать порядковыми статистиками $X_{(1)}\leqslant X_{(2)}\leqslant\ldots\leqslant X_{(r)}$ ($X_{(i)}$ — продолжительность работы изделия, отказавшего по порядку i-м). Рассмотрим класс оценок неизвестного значения $\lambda^{-1}=\mathbf{M}X_1$, имеющих вид $c_1X_{(1)}+\ldots+c_rX_{(r)}$ и удовлетворяющих условию *несмещенности*:

$$\mathbf{M}(c_1X_{(1)}+\ldots+c_rX_{(r)}) = \lambda^{-1}.$$

Найти в этом классе оценку с минимальной дисперсией, пользуясь тем, что сл. в. $Y_i=(n-i+1)(X_{(i)}-X_{(i-1)})$, $i=1, \ldots, n$, $X_{(0)}=0$ независимы и имеют показательную плотность $\lambda e^{-\lambda x}$, $x\geqslant 0$ (см. 4.17).

[Оценку можно искать в виде $U=a_1Y_1+\ldots+a_rY_r$. Имеем

$$\mathbf{M}U = \lambda^{-1}(a_1+\ldots+a_r),\ \mathbf{D}U = \lambda^{-1}(a_1^2+\ldots+a_r^2).$$

Минимум выражения $a_1^2+\ldots+a_r^2$ при условии $a_1+\ldots+a_r=1$ достигается при $a_1=\ldots=a_r=1/r$, так что оптимальная оценка имеет вид

$$U = r^{-1}\sum_{i=1}^{r}(n-i+1)(X_{(i)}-X_{(i-1)}) = r^{-1}\sum_{i=1}^{r}(n-i+1)X_{(i)} -$$

$$- r^{-1}\sum_{i=1}^{r}(n-i+1)X_{(i-1)} = r^{-1}\sum_{i=1}^{r}X_{(i)} + r^{-1}(n-r)X_{(r)}.$$

Таким образом, U представляет собой линейную функцию *полной наработки* $X_{(1)}+\ldots+X_{(r)}+(n-r)X_{(r)}$.]

Конечную последовательность сл. в. $X_1, \ldots, X_n$ бывает удобно рассматривать как вектор $\mathbf{X}=(X_1,\ldots,X_n)$ со случайными компонентами. *Математическим ожиданием* случайного вектора называют вектор, составленный из математических ожиданий его компонент:

$$\mathbf{MX}=(\mathbf{M}X_1, \ldots, \mathbf{M}X_n).$$

Ковариационной матрицей вектора X (или набора сл. в. $X_1, \ldots$ $\ldots, X_n$) называют матрицу

$$\mathbf{R_x}=\|\operatorname{cov}(X_i, X_j), \quad i, j=1, \ldots, n\|.$$

Обращение со случайными векторами упрощается, если использовать аппарат матричной алгебры. С этой целью определим математическое ожидание матрицы из случайных величин как матрицу, образованную из математических ожиданий ее компонент:

$$\mathbf{M}\|Z_{ij}, i=1, \ldots, m, j=1, \ldots, n\|=$$
$$=\|\mathbf{M}Z_{ij}, i=1, \ldots, m, j=1, \ldots, n\|.$$

Тогда ковариационную матрицу можно представить в виде

$$\mathbf{R_X}=\mathbf{M}((\mathbf{X}-\mathbf{MX})'(\mathbf{X}-\mathbf{MX})),$$

где штрих обозначает транспонирование, так что

$$(\mathbf{X}-\mathbf{MX})'(\mathbf{X}-\mathbf{MX})=\|(X_i-\mathbf{M}X_i)(X_j-\mathbf{M}X_j), \quad i, j=1, \ldots, n\|$$

есть произведение матрицы-столбца на матрицу-строку.

8.28. Задача. Пусть Z, V — случайные матрицы, A, B — постоянные матрицы. Показать, что

$$\mathbf{M}(Z+V)=\mathbf{M}Z+\mathbf{M}V, \ \mathbf{M}(AZ)=A\mathbf{M}Z, \ \mathbf{M}(ZB)=(\mathbf{M}Z)B,$$

где размерности матриц Z, V, A, B предполагаются согласованными:

$$\Big[\mathbf{M}(AZ)=\mathbf{M}\Big\|\sum_{(k)} a_{ik}Z_{kj}\Big\|=\Big\|\mathbf{M}\sum_{(k)} a_{ik}Z_{kj}\Big\|=$$
$$=\Big\|\sum_{(k)} a_{ik}\mathbf{M}Z_{kj}\Big\|=\|a_{ik}\|\cdot\|\mathbf{M}Z_{kj}\|=A\cdot\mathbf{M}Z.\Big]$$

8.29. Задача. Вывести формулу $\mathbf{R_X}=\mathbf{M}(\mathbf{X}'\mathbf{X})-(\mathbf{MX}')\cdot(\mathbf{MX})$.

$$[\mathbf{M}((\mathbf{X}-\mathbf{MX})'(\mathbf{X}-\mathbf{MX}))=\mathbf{M}(\mathbf{X}'\mathbf{X}-\mathbf{X}'\mathbf{MX}-(\mathbf{MX}')\mathbf{X}+\mathbf{MX}'\mathbf{MX})=$$
$$=\mathbf{MX}'\mathbf{X}-\mathbf{MX}'\mathbf{MX}-\mathbf{MX}'\mathbf{MX}+\mathbf{MX}'\mathbf{MX}=\mathbf{MX}'\mathbf{X}-\mathbf{MX}'\mathbf{MX}.]$$

8.30. Задача. Показать, что

$$\mathbf{R_{X+a}}=\mathbf{R_X}, \ \mathbf{R}_{\mathbf{X}A}=A'\mathbf{R_X}A,$$

где $\mathbf{a}$ — постоянный вектор, A — постоянная матрица.

$$[\mathbf{R}_{\mathbf{X}A}=\mathbf{M}((\mathbf{X}A)'\mathbf{X}A)-\mathbf{M}(\mathbf{X}A)'\mathbf{M}(\mathbf{X}A)=\mathbf{M}(A'\mathbf{X}'\mathbf{X}A)-$$
$$-\mathbf{M}(A'\mathbf{X}')\cdot\mathbf{M}(\mathbf{X}A)=A'\mathbf{M}(\mathbf{X}'\mathbf{X})A-A'\mathbf{MX}'\mathbf{MX}A=A'(\mathbf{MX}'\mathbf{X}-\mathbf{MX}'\mathbf{MX})A.]$$

8.31. Задача. Показать, что матрица ковариаций симметрична и неотрицательно определена.

[$\sum\limits_{(i,j)} c_i c_j \operatorname{cov}(X_i, X_j) = \sum\limits_{(i,j)} \operatorname{cov}(c_i X_i, c_j X_j) = \mathbf{D} \sum\limits_{(i)} c_i X_i \geqslant 0$. Отметим, что если матрица ковариаций вырождена, то найдется ненулевой вектор $(c_1, \ldots, c_n) = \mathbf{c}$, что $\mathbf{D}(\mathbf{c}\mathbf{X}') = 0$, т. е. $\mathbf{c}\mathbf{X}' = \mathbf{M}(\mathbf{c}\mathbf{X}')$ с вероятностью 1. Это означает, что совместное распределение вероятностей сл. в. $X_1, \ldots, X_n$ сосредоточено в некотором линейном подпространстве в R^n (размерности $<n$).]

8.32. Задача. Сл. в. $(X_1, \ldots, X_n) = \mathbf{X}$ имеют невырожденную матрицу ковариаций. Вывести из 8.30, что существуют ортогональная матрица C и невырожденная матрица A, такие, что случайный вектор $(Y_1, \ldots, Y_n) = \mathbf{X}C$ имеет некоррелированные компоненты, а случайный вектор $(U_1, \ldots, U_n) = \mathbf{X}A$ имеет единичную матрицу ковариаций.

[Как известно, симметричную матрицу можно привести к диагональной форме

$$C\mathbf{R}_{\mathbf{x}} C' = \Lambda$$

с помощью подходящей ортогональной матрицы C $(C'C = CC' = I)$. Если матрица $\mathbf{R}_{\mathbf{x}}$ неотрицательно определена, то на диагонали матрицы Λ стоят неотрицательные числа $\lambda_i \geqslant 0$, $i=1, \ldots, n$, если $\mathbf{R}_{\mathbf{x}}$ к тому же невырождена (положительно определена), то все $\lambda_i > 0$, $i=1, \ldots, n$. Полагая $\mathbf{Y} = \mathbf{X}C'$, находим из 8.30

$$R\mathbf{Y} = (C')'\mathbf{R}_{\mathbf{X}} C' = C\mathbf{R}_{\mathbf{X}} C' = \Lambda.$$

Домножая полученное соотношение слева и справа на диагональную матрицу $\Lambda^{-1/2}$ (с элементами $\lambda_i^{-1/2}$, $i=1, \ldots, n$, на главной диагонали), получаем

$$\Lambda^{-1/2} R\mathbf{Y}\, \Lambda^{-1/2} = I,$$

т. е. вектор $\mathbf{Y}\Lambda^{-1/2} = \mathbf{U}$ имеет единичную ковариационную матрицу. Таким образом, $\mathbf{U} = \mathbf{X}A$ и $A = \mathbf{C}'\Lambda^{-1/2}$ удовлетворяет требованиям задачи. Соотношение $\mathbf{X} = \mathbf{U}A^{-1}$ показывает, что в условиях задачи вектор $\mathbf{X}$ получается линейным преобразованием из вектора $\mathbf{U}$, компоненты которого некоррелированы и имеют единичные дисперсии.]

8.33.* Задача. Матрица ковариаций сл. в. $(X_1, \ldots, X_n) = \mathbf{X}$ имеет ранг $m < n$. Показать, что существует вектор $\mathbf{U} = (U_1, \ldots, U_m)$ с единичной ковариационной матрицей, такой, что $\mathbf{X}$ получается из $\mathbf{U}$ линейным преобразованием $\mathbf{X} = \mathbf{U}B + \mathbf{b}$ (с матрицей B ранга m).

[Как и в 8.32, находим ортогональную матрицу C, такую, что

$$C\mathbf{R}_{\mathbf{x}} C' = \Lambda,$$

где матрица Λ диагональна с числами $\lambda_1, \ldots, \lambda_m > 0$, $\lambda_{m+1} = \ldots = \lambda_n = 0$, на главной диагонали. Вектор $\mathbf{Y} = (Y_1, \ldots, Y_n) = \mathbf{X}C'$ имеет матрицу ковариаций Λ, так что $\mathbf{D}Y_k = \lambda_k$ и, следовательно, $Y_k =$

$=\mathbf{M}Y_k$, $k=m+1, \ldots, n$, с вероятностью 1. Без ограничения общности положим $\mathbf{MX}=0$, так что $\mathbf{MY}=0$ и $Y_k=0$, $k=m+1, \ldots, n$, с вероятностью 1. Таким образом, имеем

$$\mathbf{X}=\mathbf{Y}(C')^{-1}=\mathbf{Y}C=(Y_1, \ldots, Y_m, 0, \ldots, 0)C.$$

Обозначим $C_{m,n}$ матрицу, образованную из C вычеркиванием последних $n-m$ строк, а Λ_m — диагональную матрицу порядка m с элементами $\lambda_1, \ldots, \lambda_m$ на главной диагонали. Вектор $(Y_1, \ldots, Y_m)$ имеет матрицу ковариаций Λ_m, вектор $(U_1, \ldots, U_m)=(Y_1, \ldots, Y_m)\Lambda_m^{-1/2}$ имеет единичную ковариационную матрицу, так что

$$\mathbf{X}=(Y_1, \ldots, Y_m)C_{m,n}=(U_1, \ldots, U_m)\Lambda_m^{1/2}C_{m,n}=\mathbf{U}B,$$

что и требовалось установить. Отметим, что из соотношения $\mathbf{X}=\mathbf{U}B$ следует, что с вероятностью 1 вектор $\mathbf{X}$ лежит в подпространстве L в R^n, порожденном строками матрицы B (при $\mathbf{MX}=0$, а при $\mathbf{MX}\neq 0$ распределение вероятностей сосредоточено на гиперплоскости, получаемой сдвигом L на вектор $\mathbf{MX}$). Более того, если допустить, что распределение вероятностей вектора $\mathbf{X}$ сосредоточено в некотором подпространстве $L'\subseteq L$, $L'\neq L$, то, взяв какой-либо ненулевой вектор $\mathbf{a}\in L$, ортогональный к L', имеем

$$0=\mathbf{X}\mathbf{a}'=\mathbf{U}B\mathbf{a}'$$

с вероятностью 1. Откуда

$$0=\mathbf{D}(\mathbf{U}B\mathbf{a}')=\mathbf{a}B'B\mathbf{a}'=(\mathbf{a}B')(\mathbf{a}B')',$$

что противоречит предположению, что матрица B имеет ранг m, поскольку длина ненулевого вектора $\mathbf{a}B'$ оказалась равной нулю.]

8.34. Задача. Случайные векторы $\mathbf{X}=(X_1, \ldots, X_n)$ и $\mathbf{Y}=(Y_1, \ldots, Y_m)$ *некоррелированы*, т. е. $\mathrm{cov}(X_i, Y_j)=0$, $i=1, \ldots, n$, $j=1, \ldots, m$. Показать, что векторы $\mathbf{X}A$ и $\mathbf{Y}B$, где A, B — некоторые постоянные матрицы, также некоррелированы.

[Для любых векторов $\mathbf{a}=(a_1, \ldots, a_n)$, $\mathbf{b}=(b_1, \ldots, b_m)$ имеем

$$\mathrm{cov}(\mathbf{X}\mathbf{a}', \mathbf{Y}\mathbf{b}')=\sum_{(i,j)} a_i b_j \,\mathrm{cov}(X_i, Y_j)=0.]$$

8.35.* Задача. Пусть случайный вектор $\mathbf{X}=(X_1, \ldots, X_n)$ имеет ковариационную матрицу ранга $0<m<n$. Показать, что для некоторых $1\leqslant i_1<\ldots<i_m\leqslant n$ сл. в. $X_{i_1}, \ldots, X_{i_m}$ имеют невырожденную ковариационную матрицу, а остальные координаты X_i вектора $\mathbf{X}$ линейно выражаются (с вероятностью 1) через $X_{i_1}, \ldots, X_{i_m}$.

[Переставим координаты вектора $\mathbf{X}$ таким образом, чтобы угловой минор порядка m матрицы $\mathbf{R}_X$ был положителен. Покажем, что в этом случае требованиям задачи удовлетворяет набор индексов $i_k=k$, $k=1, \ldots, m$. Во-первых, матрица ковариаций вектора $(X_1, \ldots, X_m)$ получается из $\mathbf{R}_X$ вычеркиванием последних $n-$

—m строк и столбцов и, следовательно, невырождена. Далее запишем

$$\mathbf{D}(a_1X_1+\ldots+a_mX_m+a_{m+i}X_{m+i})=\mathbf{a}\mathbf{R}_X\mathbf{a}',$$

$$\mathbf{a}=(a_1, \ldots, a_m, 0, \ldots, 0, a_{m+i}, 0, \ldots, 0)\in R^n.$$

Предположив, что для любого вектора $\mathbf{a}\neq 0$ указанного вида $\mathbf{a}\mathbf{R}_x\mathbf{a}'\neq 0$, мы бы имели положительно определенную квадратичную форму, составленную из подматрицы $\mathbf{R}_x$ размера $(m+1)\times\times(m+1)$, чего не может быть, так как $\mathbf{R}_X$ имеет ранг m. Следовательно, для некоторого $\mathbf{a}=(a_1, \ldots, a_m, 0, \ldots, 0, a_{m+i}, 0, \ldots, 0)\neq\neq 0$ имеем $\mathbf{a}\mathbf{R}_x\mathbf{a}'=0$ и $a_1X_1+\ldots+a_mX_m+a_{m+1}X_{m+1}=\mathrm{const}$ с вероятностью 1. Остается заметить, что $a_{m+i}\neq 0$, так как иначе сл. в. $X_1, \ldots, X_m$ были бы линейно зависимы и их матрица ковариаций оказалась бы вырожденной.]

8.36.* Задача. Допустим, что математическое ожидание и ковариационная матрица случайного вектора $\mathbf{X}=(X_1, \ldots, X_n)$ известны. Построить линейную оценку $\widehat{X}_n=a_1X_1+\ldots+a_{n-1}X_{n-1}+ +b$ сл. в. X_n наилучшую в среднеквадратическом смысле: $\mathbf{M}(X_n-\widehat{X}_n)^2=\min$, предполагая, что $\mathbf{R}_x$ невырождена (ср. 8.22).

[Ввиду 8.3 минимум по b среднеквадратического расстояния $(\mathbf{M}(X_n-\widehat{X}_n)^2)^{1/2}$ достигается при $b=\mathbf{M}X_n-a_1\mathbf{M}X_1-\ldots-a_{n-1}\mathbf{M}X_{n-1}$, так что далее предположим, что $\mathbf{M}X_i=0$, $i=1,\ldots,n$. Дифференцируя по $a_1,\ldots,a_{n-1}$ выражение

$$\mathbf{M}(X_n-\widehat{X}_n)^2=\langle X_n-a_1X_1-\ldots-a_{n-1}X_{n-1},\ X_n-a_1X_1-\ldots-a_{n-1}X_{n-1}\rangle=$$

$$=\sum_{(i,j)}a_ia_j\langle X_i, X_j\rangle,\ a_n=-1,\ \langle X_i, X_j\rangle\equiv\mathbf{M}X_iX_j$$

и приравнивая нулю производные (либо сразу записывая условия ортогональности вектора $X_n-\widehat{X}_n$ к каждому из векторов $X_1, \ldots, X_{n-1}$), получаем

$$\langle X_n, X_i\rangle-a_1\langle X_1, X_i\rangle-\ldots-a_{n-1}\langle X_{n-1}, X_i\rangle=0,\quad i=1, \ldots, n-1.$$

Пусть $\mathbf{R}_\mathbf{X}=\|r_{ij},\ i, j=1, \ldots, n\|$, $Q=\|q_{ij},\ i, j=1, \ldots, n\|=\mathbf{R}_\mathbf{X}^{-1}$, Δ — определитель матрицы $\mathbf{R}_x$, A_{ij} — алгебраическое дополнение элемента r_{ij} в $\mathbf{R}_x$. Решая линейную систему по правилу Крамера, находим

$$a_i=\begin{vmatrix} r_{11} & \cdots & r_{1\,i-1} & r_{1n} & r_{1\,i+1} & \cdots & r_{1\,n-1}\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ r_{n-1\,1} & \cdots & r_{n-1\,i-1} & r_{n-1\,n} & r_{n-1\,i+1} & \cdots & r_{n-1\,n-1}\end{vmatrix}\cdot\begin{vmatrix} r_{11} & \cdots & r_{1\,n-1}\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\\ r_{n-1\,1} & \cdots & r_{n-1\,n-1}\end{vmatrix}^{-1}=$$

$$=-A_{ni}/A_{nn}=-(A_{ni}\Delta^{-1})/(A_{nn}\Delta^{-1})=-q_{ni}/q_{nn},$$

$$\widehat{X}_n=-(q_{n1}/q_{nn})X_1-\ldots-(q_{n\,n-1}/q_{nn})X_{n-1}.]$$

§ 9.

УСЛОВНОЕ МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ

Пусть $(\Omega, \mathscr{B}, \mathbf{P})$ — основное вероятностное пространство. Для любого события A, $\mathbf{P}(A)>0$, функция множества

$$\mathbf{P}_A(B)=\mathbf{P}(AB)/\mathbf{P}(A)$$

обладает всеми свойствами вероятностной меры ($\mathbf{P}_A(\Omega)=1$, $\mathbf{P}_A(\bar{B})=1-\mathbf{P}_A(B)$, $\mathbf{P}_A(B)$ счетно-аддитивна) и называется условной вероятностной мерой при условии события A. Вероятностное пространство $(\Omega, \mathscr{B}, \mathbf{P}_A)$ рассматриваем как вспомогательное по отношению к основному пространству $(\Omega, \mathscr{B}, \mathbf{P})$. Операцию математического ожидания в $(\Omega, \mathscr{B}, \mathbf{P}_A)$ обозначим $\mathbf{M}_A$. Напомним, что действительная функция $X(\omega)$ — случайная величина, если для любого промежутка $I_{a,b}$ множество $\{\omega: X(\omega)\in I_{a,b}\}\in\mathscr{B}$. Это означает, что сл. в. $X(\omega)$, заданная на основном вероятностном пространстве, также является случайной величиной на вспомогательном пространстве $(\Omega, \mathscr{B}, \mathbf{P}_A)$. Математическое ожидание $\mathbf{M}_A X$ сл. в. $X(\omega)$, рассматриваемой на $(\Omega, \mathscr{B}, \mathbf{P}_A)$, называется *условным математическим ожиданием сл. в.* $X(\omega)$ *на* $(\Omega, \mathscr{B}, \mathbf{P})$ *при условии события* A. Наряду с $\mathbf{M}_A X$ употребляется обозначение $\mathbf{M}(X|A)$. Поскольку условное математическое ожидание определяется через обычное, оно и обладает всеми свойствами математического ожидания. Обычно интерес представляет не отдельная условная вероятностная мера $\mathbf{P}_A$, а совокупность условных вероятностей $\mathbf{P}_{A_i}$, $i=1, 2, \ldots,$ где $\{A_i, i=1, 2, \ldots\}$ — разбиение пространства Ω (или, иначе, полная группа событий). Соответственно имеем совокупность условных математических ожиданий $\mathbf{M}(X|A_i)$ сл. в. X.

9.1. Задача. Вывести формулу $\mathbf{M}(X|A)=\mathbf{M}(XI_A)/\mathbf{P}(A)$ в предположении, что $\mathbf{M}(XI_A)$ конечно.

$$[\mathbf{M}(X|A)=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty} kn^{-1}\mathbf{P}_A(\omega: kn^{-1}\leqslant X(\omega)<(k+1)n^{-1})=$$

$$=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty} kn^{-1}\mathbf{P}(\{\omega: kn^{-1}\leqslant X(\omega)<(k+1)n^{-1}\}\bigcap A)/\mathbf{P}(A)=$$

$$=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty} kn^{-1}\mathbf{P}(\omega: kn^{-1}\leqslant X(\omega)\,I_A(\omega)<(k+1)n^{-1})/\mathbf{P}(A),$$

так как при $k\neq 0$

$$\{\omega: kn^{-1}\leqslant X(\omega)<(k+1)n^{-1}\}\bigcap A=$$

$$=\{\omega: kn^{-1}\leqslant X(\omega)\,I_A(\omega)<(k+1)n^{-1}\}.]$$

9.2. Задача. Пусть события $A_1, \ldots, A_n$ образуют (конечное) разбиение пространства Ω. Вывести формулу

$$\mathbf{M}X = \sum_{i=1}^{n} \mathbf{M}(X|A_i)\,\mathbf{P}(A_i),$$

предполагая, что $\mathbf{M}X$ конечно.

[Так как $|XI_A| \leqslant |X|$, то из конечности $\mathbf{M}X$ следует конечность $\mathbf{M}(XI_A)$ для любого события A. Беря математическое ожидание от обеих частей равенства

$$X(\omega) = \sum_{i=1}^{n} X(\omega)\, I_{A_i}(\omega)$$

и применяя 9.1, получаем требуемый результат.]

9.3. Задача. Пусть события $A_1, A_2, \ldots$ образуют (счетное) разбиение пространства Ω. Вывести формулу

$$\mathbf{M}X = \sum_{i=1}^{\infty} \mathbf{M}(XI_{A_i}) = \sum_{i=1}^{\infty} \mathbf{M}(X|A_i)\,\mathbf{P}(A_i)$$

сначала для дискретной сл. в. X с конечным $\mathbf{M}X$, а затем для произвольной.

[Для дискретной сл. в. X имеем

$$\mathbf{M}X = \sum_{\{x\}} x\mathbf{P}(\omega : X(\omega) = x) = \sum_{\{x\}} x \sum_{i=1}^{\infty} \mathbf{P}(\{\omega : X(\omega) = x\} \cap A_i) =$$

$$= \sum_{i=1}^{\infty} \sum_{\{x\}} x\mathbf{P}(\omega : X(\omega)\, I_{A_i}(\omega) = x) = \sum_{i=1}^{\infty} \mathbf{M}(XI_{A_i}).$$

Для произвольной сл. в. X образуем последовательность дискретных случайных величин $X^{(n)}(\omega) = kn^{-1}$ при $kn^{-1} < X(\omega) \leqslant (k+1)n^{-1}$, так что

$$\mathbf{M}X = \lim_{n\to\infty} \mathbf{M}X^{(n)} = \lim_{n\to\infty} \sum_{i=1}^{\infty} \mathbf{M}(X^{(n)} I_{A_i}) = \lim_{n\to\infty} \sum_{i=1}^{\infty} \mathbf{M}(X^{(n)}|A_i)\,\mathbf{P}(A_i).$$

Так как $|X^{(n)}(\omega) - X(\omega)| \leqslant n^{-1}$, то $|\mathbf{M}(X^{(n)}|A_i) - \mathbf{M}(X|A_i)| \leqslant n^{-1}$ и

$$\left| \sum_{i=1}^{\infty} \mathbf{M}(X^{(n)}|A_i)\,\mathbf{P}(A_i) - \sum_{i=1}^{\infty} \mathbf{M}(X|A_i)\,\mathbf{P}(A_i) \right| \leqslant n^{-1}.]$$

9.4. Пример (оптимальная нелинейная оценка). Предположим, что наблюдается дискретная сл. в. Y (или набор таких величин), а сл. в. X с $\mathbf{M}X^2 < \infty$ ненаблюдаема. Ищем оценку сл. в. X по сл. в. Y в виде произвольной функции $g(Y)$, такой, что

$\mathbf{M}g(Y)^2<\infty$. Оценку $\widehat{X}=\widehat{g}(Y)$ назовем *оптимальной в среднеквадратичном*, если

$$\mathbf{M}(X-\widehat{X})^2=\min \mathbf{M}(X-g(Y))^2$$

по всевозможным функциям $g(y)$.

Применяя 9.3, имеем

$$\mathbf{M}(X-g(Y))^2=\sum_{\{y\}} \mathbf{M}((X-g(Y))^2 I_{\{\omega:Y(\omega)=y\}})=$$

$$=\sum_{\{y\}} \mathbf{M}((X-g(y))^2 I_{\{\omega:Y(\omega)=y\}})=$$

$$=\sum_{\{y\}} \mathbf{M}((X-g(y))^2 \mid \{\omega: Y(\omega)=y\})\cdot \mathbf{P}(\omega: Y(\omega)=y).$$

Поскольку для любого события A

$$\min_{\{a\}} \mathbf{M}_A(X-a)^2=\mathbf{M}_A(X-\mathbf{M}_A X)^2,$$

то при каждом значении y сл. в. Y оптимальная оценка имеет вид

$$\hat{g}(y)=\mathbf{M}(X\mid\{\omega: Y(\omega)=y\}) \quad (\mathbf{P}(\omega): Y(\omega)=y)>0).$$

Итак, оптимальная нелинейная оценка дается функцией $\hat{g}(y)$, представляющей собой при каждом y условное математическое ожидание сл. в. X при условии события $\{\omega: Y(\omega)=y\}$. Оценка $\hat{g}(Y)$ — случайная величина, функция от Y. Ее называют *условным математическим ожиданием*, или *условным средним сл. в. X при условии сл. в. Y*, и обозначают

$$\hat{g}(Y)=\mathbf{M}(X\mid Y).$$

(С точки зрения геометрии гильбертова пространства множество сл. в. $g(Y)$, $\mathbf{M}g(Y)^2<\infty$, образует замкнутое линейное подпространство в пространстве H всех случайных величин на $(\Omega, \mathscr{B}, \mathbf{P})$ с конечным математическим ожиданием квадрата, а $\mathbf{M}(X\mid Y)$ представляет собой проекцию вектора X на это подпространство.)

9.5. Задача. Сл. в. X, Y независимы и имеют биномиальные распределения с параметрами n, p и m, p соответственно. Найти $\mathbf{M}(X\mid X+Y)$ (условное распределение см. в 5.4)

[Так как (см. 5.4)

$$p_{X\mid(X+Y)}(k\mid l)=C_n^k C_m^{l-k}/C^l_{n+m},$$

то (см. 7.2)

$$\mathbf{M}(X\mid\{\omega: X(\omega)+Y(\omega)=l\})=ln/(n+m),$$

$$\mathbf{M}(X\mid X+Y)=(X+Y)n/(n+m).$$

Полученный результат можно интерпретировать следующим образом. Допустим, что известно общее число $X+Y$ успехов в $n+$ $+m$ испытаниях Бернулли. Тогда наилучшая среднеквадратичная оценка числа X успехов в первых n испытаниях — величина $(X+Y)n/(n+m)$, т. е. пропорциональная доля общего числа успехов.]

9.6. Задача. Сл. в. N с целыми неотрицательными значениями не зависит от последовательности сл. в. $X_1, X_2, \ldots, S_N=X_1+$ $+\ldots+X_N$ ($S_N=0$ при $N=0$). Предполагая, что $\mathbf{M}X_i=a$, $i=1, 2, \ldots$ $\ldots$, $\mathbf{M}N$ конечны, найти $\mathbf{M}(S_N|N)$.

[Воспользовавшись 9.1, запишем

$$\mathbf{M}(S_N|\{\omega : N(\omega)=n\})=\mathbf{M}(S_N I_{\{\omega:N(\omega)=n\}})/\mathbf{P}(\omega : N(\omega)=n).$$

Далее имеем

$$\mathbf{M}(S_N I_{\{\omega:N(\omega)=n\}})=\mathbf{M}(S_n I_{\{\omega:N(\omega)=n\}})=\mathbf{M}S_n\,\mathbf{M}I_{\{\omega:N(\omega)=n\}}=$$
$$=na\mathbf{P}(\omega : N(\omega)=n),\quad \mathbf{M}(S_N|N)=Na.]$$

Остановимся подробнее на определении условного математического ожидания. Пусть на вероятностном пространстве $(\Omega, \mathscr{B}, \mathbf{P})$ заданы сл. в. X, Y, причем сл. в. Y дискретна (либо в качестве Y возьмем набор дискретных сл. в. $(Y_1, \ldots, Y_n)$). Введем разбиение $\{A_y\}$, $A_y=\{\omega: Y(\omega)=y\}$, порожденное сл. в. Y (или набором случайных величин $\mathbf{Y}=(Y_1, \ldots, Y_n)$). *Определим* математическое ожидание $\mathbf{M}(X|Y)$ формулой

$$\mathbf{M}(X|Y)=\Sigma_{\{y\}}\mathbf{M}(X|A_y)\cdot I_{A_y}\qquad (\mathbf{P}(A_y)>0),$$

где I_A — индикатор события A. Для существования $\mathbf{M}(X|Y)$ необходимо и достаточно, чтобы при любом y было конечно $\mathbf{M}(X|A_y)$. Это обеспечивается конечностью $\mathbf{M}X$ (см. 9.3). Последнее условие принимается без оговорок всякий раз, когда речь будет идти об условном математическом ожидании сл. в. X. Как видно из определения, $\mathbf{M}(X|Y)$ — функция на Ω, принимающая постоянные значения на элементах разбиения $\{A_y\}: \mathbf{M}(X|Y)=$ $=\mathbf{M}(X|A_y)$ при $\omega\in A_y$. Можно сказать, что $\mathbf{M}(X|Y)$ получается из X усреднением: значения $X(\omega)$ при $\omega\in A_y$ усредняются по условной мере $\mathbf{P}_{A_y}$, и усредненное значение принимается за значение сл. в. $\mathbf{M}(X|Y)$ при $\omega\in A_y$.

Роль сл. в. Y при определении $\mathbf{M}(X|Y)$ сводится к разбиению пространства Ω, порождаемому этой случайной величиной. Поэтому наряду с понятием условного математического ожидания относительно случайной величины можно говорить об *условном математическом ожидании при условии произвольного* (не более чем счетного) *разбиения* $\{A_i, i=1, 2, \ldots\}$. Обозначим $\mathscr{A}$ систему всех подмножеств из Ω, которые являются объединениями элементов разбиения $\{A_i, i=1, 2, \ldots\}$ (и включающую пустое подмножество). Система $\mathscr{A}$, очевидно, является алгеброй множеств, и, кроме того, она замкнута относительно образования счетного объединения своих элементов. Любой класс подмножеств с указанными свойствами называется *σ-алгеброй множеств*. По ряду причин удобнее говорить об условном математическом ожидании сл. в. относительно σ-алгебры $\mathscr{A}=\sigma(\{A_i,\ i=1,\ 2,\ \ldots\})$, *порожденной* разбиением $\{A_i,\ i=1,\ 2,\ldots\}$, нежели об условном математическом ожидании относительно самого разбиения. *Определим*

$$\mathbf{M}(X\,|\,\mathscr{A})=\sum_{(i)}\mathbf{M}(X\,|\,A_i)\,I_{A_i}.$$

9.7. Задача. Установить равенства

$$\mathbf{M}(aX+b\,|\,\mathscr{A})=a\mathbf{M}(X\,|\,\mathscr{A})+b,$$

$$\mathbf{M}(X_1+X_2\,|\,\mathscr{A})=\mathbf{M}(X_1\,|\,\mathscr{A})+\mathbf{M}(X_2\,|\,\mathscr{A}).$$

[Вытекают из определения $\mathbf{M}(X\,|\,\mathscr{A})$ и того, что $\mathbf{M}(X\,|\,A)=\mathbf{M}_AX$ — обычное математическое ожидание на вспомогательном вероятностном пространстве $(\Omega,\ \mathscr{B},\ \mathbf{P}_A)$.]

9.8. Задача. Установить формулу $\mathbf{M}(\mathbf{M}(X\,|\,\mathscr{A}))=\mathbf{M}X$.
[Совпадает с формулой 9.2.]

9.9. Задача. Пусть сл. в. Y постоянна на элементах разбиения $\{A_i,\ i=1,\ 2,\ \ldots\}$, $\mathscr{A}=\sigma(\{A_i,\ i=1,\ 2,\ \ldots\})$. Показать, что $\mathbf{M}(XY\,|\,\mathscr{A})=Y\mathbf{M}(X\,|\,\mathscr{A})$, $\mathbf{M}(Y\,|\,\mathscr{A})=Y$.
[При вычислении $\mathbf{M}(XY\,|\,A_i)$ заметим, что сл. в. Y, рассматриваемая на вспомогательном вероятностном пространстве $(\Omega,\ \mathscr{B},\ \mathbf{P}_{A_i})$, постоянна и потому выносится за знак математического ожидания $\mathbf{M}_{A_i}$. Таким образом,

$$\mathbf{M}(XY\,|\,\mathscr{A})=\sum_{(i)}Y\mathbf{M}(X\,|\,A_i)\,I_{A_i}.$$

Учитывая, что для каждого ω ровно для одного значения i имеем $I_{A_i}(\omega)=1$, а для остальных $j\neq i$ $I_{A_j}(\omega)=0$, можно вынести Y за знак суммы и получить

$$\mathbf{M}(XY\,|\,\mathscr{A})=Y\sum_{(i)}\mathbf{M}(X\,|\,A_i)\,I_{A_i}=Y\mathbf{M}(X\,|\,\mathscr{A}).$$

Рассматривая разбиение, порожденное произвольной дискретной сл. в. Z, перепишем полученное свойство в виде

$$\mathbf{M}(Xh(Z)\,|\,Z)=h(Z)\,\mathbf{M}(X\,|\,Z)$$

для любой функции $h(z)$.]

9.10. Задача. Если сл. в. X и Y независимы, то $\mathbf{M}(X\,|\,Y)=$ $=\mathbf{M}(X)$.
[Используя 9.1, **имеем**

$$\mathbf{M}(X\,|\,Y)=\sum_{(y)}\mathbf{M}(XI_{A_y})\,\mathbf{P}(A_y)^{-1}I_{A_y}=\sum_{(y)}(\mathbf{M}X)\cdot I_{A_y}=\mathbf{M}X.$$

В случае, когда $\mathbf{M}X^2<\infty$, получаем, что наилучшая среднеквадратичная оценка сл. в. X по независящей от нее сл. в. Y есть постоянная, равная $\mathbf{M}X$, что вполне понятно.]

9.11. Задача. Показать, что

$$\mathbf{M}(\mathbf{M}(X \mid Y, Z) \mid Y) = \mathbf{M}(X \mid Y).$$

$$[\mathbf{M}(\mathbf{M}(X \mid Y, Z) \mid \{\omega : Y(\omega) = y\}) = \mathbf{M}(\mathbf{M}(X \mid Y, Z) \cdot I_{\{\omega: Y(\omega)=y\}}) \times$$

$$\times \cdot \mathbf{P}(\omega : Y(\omega) = y)^{-1} = \mathbf{M}\mathbf{M}(X I_{\{\omega: Y(\omega)=y\}} \mid Y, Z) \cdot \mathbf{P}(\omega : Y(\omega) = y)^{-1} =$$

$$= \mathbf{M}(X I_{\{\omega: Y(\omega)=y\}}) \mathbf{P}(\omega : Y(\omega) = y)^{-1} = \mathbf{M}(X \mid \{\omega : Y(\omega) = y\}),$$

где мы последовательно использовали 9.1, 9.9, 9.8 и 9.1.

(На языке геометрии гильбертова пространства при $\mathbf{M}X^2 < \infty$ $\mathbf{M}(X \mid Y, Z)$ является проекцией вектора X на подпространство $L_{y,z}$ сл. в. $g(Y, Z)$, а $\mathbf{M}(X \mid Y)$ — проекция X на подпространство L_y сл. в. $h(Y)$. Очевидно, $L_y \subseteq L_{y,z}$, и потому проекцию X на L_y можно получить, сначала спроектировав X на $L_{y,z}$, а затем вектор проекции еще раз спроектировать на L_y.)]

9.12. Задача. Пусть $\mathscr{A}^{(1)}$, $\mathscr{A}^{(2)}$ — σ-алгебры подмножеств из Ω, порожденные (счетными) разбиениями $\{A_i^{(1)}\}$, $\{A_j^{(2)}\}$ соответственно, причем $\mathscr{A}^{(1)} \subseteq \mathscr{A}^{(2)}$. Показать, что

$$\mathbf{M}(\mathbf{M}(X \mid \mathscr{A}^{(2)}) \mid \mathscr{A}^{(1)}) = \mathbf{M}(X \mid \mathscr{A}^{(1)}).$$

[По условию $A_i^{(1)} \in \mathscr{A}^{(2)}$, так что $A_i^{(1)}$ представимо в виде объединения элементов из $\mathscr{A}^{(2)}$, и потому индикаторная сл. в. $I_{A_i^{(1)}}$ постоянна на элементах разбиения $\{A_j^{(2)}\}$. Дальнейшая выкладка есть повторение 9.11:

$$\mathbf{M}(\mathbf{M}(X \mid \mathscr{A}^{(2)}) \mid A_i^{(1)}) = \mathbf{M}(\mathbf{M}(X \mid \mathscr{A}^{(2)}) I_{A_i^{(1)}}) \cdot \mathbf{P}(A)_i^{(1)})^{-1} =$$

$$= \mathbf{M}(\mathbf{M}(X I_{A_i^{(1)}} \mid \mathscr{A}^{(2)})) \cdot \mathbf{P}(A_i^{(1)})^{-1} = \mathbf{M}(X I_{A_i^{(1)}}) \mathbf{P}(A_i^{(1)})^{-1} = \mathbf{M}(X \mid A_i^{(1)}).]$$

Перепишем формулу, определяющую $\mathbf{M}(X \mid Y)$ для *дискретных* сл. в. X, Y (или $\mathbf{Y} = (Y_1, \ldots, Y_n)$), выразив $\mathbf{M}(X \mid \{\omega : Y(\omega) = y\})$ через условное распределение $p_{X \mid Y}(x \mid y)$:

$$\mathbf{M}(X \mid Y) = \sum_{\{y\}} I_{A_y} \sum_{\{x\}} x p_{X \mid Y}(x \mid y) = \sum_{\{x\}} x \sum_{\{y\}} I_{A_y} p_{X \mid Y}(x \mid y).$$

Учитывая, что при каждом ω

$$I_{A_y}(\omega) p_{X \mid Y}(x \mid y) = I_{A_y}(\omega) p_{X \mid Y}(x \mid Y(\omega)),$$

получаем, вынося из-под внутренней суммы множитель $p_{X \mid Y}(x \mid Y)$,

$$\mathbf{M}(X \mid Y) = \sum_{\{x\}} x p_{X \mid Y}(x \mid Y) \sum_{\{y\}} I_{A_y}$$

или

$$\mathbf{M}(X \mid Y) = \sum_{\{x\}} x p_{X \mid Y}(x \mid Y).$$

Выражение $p_{X|Y}(x|Y)$ при каждом фиксированном x — случайная величина, действительная функция от Y. Рассматривая $p_{X|Y}(x|Y)$ как функцию от x, зависящую от случайного параметра Y, получаем случайное распределение вероятностей на множестве $\mathscr{X}$ значений сл. в. X, т. е. отображение, ставящее в соответствие каждому значению y сл. в Y распределение вероятностей $p_{X|Y}(x|y)$ на $\mathscr{X}$.

Определение условного математического ожидания в вышеуказанной форме непосредственно обобщается на сл. в X, Y (или $\mathbf{Y}=(Y_1, \ldots, Y_n)$), обладающие совместной плотностью распределения вероятностей. Именно, положим

$$\mathbf{M}(X|Y)=\int_{-\infty}^{\infty} x f_{X|Y}(x|Y)\,dx,$$

где

$$f_{X|Y}(x|y)=f_{X,Y}(x,\ y)/f_Y(y)$$

— условная плотность сл. в. X при условии Y. Введем также понятие условного математического ожидания сл. в. X при фиксированном значении сл. в. Y, полагая

$$\mathbf{M}_y X=\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx.$$

В отличие от дискретного случая, когда событие $\{\omega: Y(\omega)=y\}$ имело положительную вероятность и можно было определить $\mathbf{M}(X|\{\omega: Y(\omega)=y\})$, в непрерывном случае аналогичное понятие не имеет столь прозрачного смысла. Однако и здесь $\mathbf{M}(X|\mathbf{Y})$ — результат подстановки сл. в. Y вместо аргумента y действительной функции $g(y)=\mathbf{M}_y X$.

Условные плотности $f_{X|Y}(x|y)$ задают на R семейство распределений вероятностей $\mathbf{P}_y$, а $\mathbf{M}_y$ представляет собой операцию математического ожидания в вероятностной модели $(R,\ \mathscr{B},\ \mathbf{P}_y)$ (где $\mathscr{B}$ — семейство борелевских множеств на прямой). Отсюда можно получить свойства условного математического ожидания $\mathbf{M}(X|Y)$ так же, как это было сделано в дискретном случае. Имеются, однако, некоторые технические затруднения. Во-первых, интеграл Римана в ряде отношений допускает меньшую свободу действий, чем суммы и ряды. Поэтому, оставаясь в рамках теории интегрирования по Риману, приходится требовать от функций $f_{X,Y}(x,\ y)$, $f_X(x)$, $f_{X|Y}(x|y)$ и т. д. дополнительных условий регулярности, чтобы обеспечить выполнение свойств $\mathbf{M}(X|Y)$. Для более широкого понятия интеграла (по Лебегу), о чем будет речь ниже, отмеченные затруднения отпадают. Мы ограничимся выводом свойств $\mathbf{M}(X|Y)$, не оговаривая требований, которые обеспечивают законность производимых манипуляций с интегралами.

Другое замечание касается формулы вычисления математического ожидания функции от случайной величины

$$\mathbf{M}g(X)=\int_{-\infty}^{\infty} g(x)f_X(x)\,dx.$$

Частный случай $g(x)=x^r$, $r>0$, был рассмотрен в 7.29 (отметим, что приведенное там рассуждение годится для любой монотонной функции). В общем виде формула будет установлена в § 10.

9.13. Задача. Показать, что $\mathbf{M}(\mathbf{M}(X\mid Y))=\mathbf{M}X$.

$$\Big[\mathbf{M}(\mathbf{M}(X\mid Y))=\int_{-\infty}^{\infty}\mathbf{M}_yX\cdot f_Y(y)\,dy=\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty}xf_{X\mid Y}(x\mid y)\,dx\Big)f_Y(y)\,dy=$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{X,Y}(x,\ y)\,dx\,dy=\int_{-\infty}^{\infty}x\Big(\int_{-\infty}^{\infty}f_{X,Y}(x,\ y)\,dy\Big)dx=$$

$$=\int_{-\infty}^{\infty}xf_X(x)\,dx=\mathbf{M}X.\Big]$$

9.14. Задача. Показать, что

$$\mathbf{M}(aX+b\mid Y)=a\mathbf{M}(X\mid Y)+b,\ \mathbf{M}(X_1+X_2\mid Y)=\mathbf{M}(X_1\mid Y)+\mathbf{M}(X_2\mid Y).$$

[Вытекает из равенств $\mathbf{M}_y(aX+b)=a\mathbf{M}_yX+b$, $\mathbf{M}_y(X_1+X_2)=\mathbf{M}_yX_1+\mathbf{M}_yX_2$.]

9.15. Задача. Показать, что $\mathbf{M}(Xh(Y)\mid Y)=h(Y)\mathbf{M}(X\mid Y)$, $\mathbf{M}(h(Y)\mid Y)=h(Y)$.

[Вытекает из равенства $\mathbf{M}_y(Xh(y))=h(y)\mathbf{M}_yX$.]

9.16. Задача. Если сл. в. X, Y независимы, то $\mathbf{M}(X\mid Y)=\mathbf{M}X$.

[При независимости X и Y имеем $f_{X\mid Y}(x\mid y)=f_X(x)$, и операции $\mathbf{M}_y$ и $\mathbf{M}$ совпадают. (Отметим, что было бы точнее сказать, что из независимости $\mathbf{X}$ и $\mathbf{Y}$ вытекает совпадение мер $\mathbf{P}_y$ и $\mathbf{P}$ на R для множества значений y сл. в. Y, имеющего вероятность 1.)]

9.17. Задача. Показать, что при $\mathbf{M}X^2<\infty$ минимум среднеквадратичного расстояния $(\mathbf{M}(X-g(Y))^2)^{1/2}$ достигается на функции $\hat{g}(y)=\mathbf{M}_yX$.

$$\Big[\mathbf{M}(X-g(Y))^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-g(y))^2f_{X,Y}(x,\ y)\,dx\,dy=$$

$$=\int_{-\infty}^{\infty}f_Y(y)\Big(\int_{-\infty}^{\infty}(x-g(y))^2f_{X\mid Y}(x\mid y)\,dx\Big)dy=\int_{-\infty}^{\infty}f_Y(y)\mathbf{M}_y(X-g(y))^2\,dx,$$

но математическое ожидание под знаком интеграла достигает минимума при $g(y)=\mathbf{M}_yX$.]

9.18. Пример (многомерное нормальное распределение). Распределение вероятностей в R^2, задаваемое плотностью вида

$$f(x_1, x_2) = c \exp\left(-\frac{1}{2}(q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2)\right),$$

где матрица $\|q_{ij}\|$ — симметричная положительно определенная, c — нормирующая константа, называется *двумерным невырожденным нормальным распределением* с нулевым средним.

Ортогональным преобразованием (поворот и симметрия) перейдем к новой системе координат (y_1, y_2):

$$x_1 = c_{11}y_1 + c_{12}y_2, \quad x_2 = c_{21}y_1 + c_{22}y_2,$$

такой, чтобы квадратичная форма $q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2$ привелась к каноническому виду $\lambda_1 y_1^2 + \lambda_2 y_2^2$, $\lambda_1, \lambda_2 > 0$. Таким образом, в новой системе координат плотность распределения вероятностной массы имеет вид

$$g(y_1, y_2) = c \exp(-{}^1/_2(\lambda_1 y_1^2 + \lambda_2 y_2^2)).$$

Интегрируя $g(y_1, y_2)$ по всему пространству и учитывая, что функция $(\lambda/2\pi)^{1/2}\exp(-(1/2)\lambda y^2)$ — плотность нормального распределения со средним 0 и дисперсией λ^{-1}, находим $c = (2\pi\sigma_1\sigma_2)^{-1}$, $\sigma_1^2 = \lambda_1^{-1}$, $\sigma_2^2 = \lambda_2^{-1}$ и убеждаемся, что $g(y_1, y_2)$ — плотность вектора (Y_1, Y_2) из независимых нормально распределенных сл. в. Y_1, Y_2, $\mathbf{M}Y_1 = \mathbf{M}Y_2 = 0$, $\mathbf{D}Y_1 = \sigma_1^2$, $\mathbf{D}Y_2 = \sigma_2^2$.

Итак, нормальное распределение на плоскости в подходящей системе координат представляет собой совместное распределение вероятностей пары независимых случайных величин, каждая из которых распределена нормально. Сл. в.

$$X_1 = c_{11}Y_1 + c_{12}Y_2, \qquad X_2 = c_{21}Y_1 + c_{22}Y_2$$

имеют совместную плотность $f(x_1, x_2)$ (см. 3.26). Пару сл. в. $(X_1, X_2) = \mathbf{X}$ называют *двумерным нормальным случайным вектором*. Из указанного представления вытекает (см. 4.20), что X_1, X_1 и любая нетривиальная линейная комбинация $a_1X_1 + a_2X_2$ имеют нормальные распределения.

Отсюда же легко установить вероятностный смысл матрицы $\|q_{ij}\|$ квадратичной формы в определении $f(x_1, x_2)$. Так как

$$\mathbf{X} = \mathbf{Y}C', \qquad C = \|c_{ij}\|,$$

то (см. 8.30)

$$\mathbf{R_X} = C\mathbf{R_Y}C', \quad \mathbf{R_X^{-1}} = C\mathbf{R_Y^{-1}}C', \quad \mathbf{R_Y} = \begin{Vmatrix} \lambda_1^{-1} & 0 \\ 0 & \lambda_2^{-1} \end{Vmatrix}.$$

Но

$$q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2 = (x_1, x_2)\,\|q_{ij}\|\,(x_1, x_2)' =$$

$$= (y_1, y_2)C'\,\|q_{ij}\|\,C(y_1, y_2)' = (y_1, y_2)\mathbf{R_Y^{-1}}(y_1, y_2)' =$$

$$=(x_1,\ x_2)CR_Y^{-1}C'(x_1,\ x_2)'=(x_1,\ x_2)R_X^{-1}(x_1,\ x_2)',$$

так что матрица $\|q_{ij}\|$ — обратная к матрице ковариаций вектора $\mathbf{X}=(X_1,\ X_2)$.

Наконец, заметим, что нормирующая постоянная c равна

$$c=(2\pi\sigma_1\sigma_2)^{-1}=(2\pi)^{-1}(\lambda_1\lambda_2)^{1/2}=(2\pi)^{-1}(\det \mathbf{R_Y})^{-1/2}=(2\pi)^{-1}(\det \mathbf{R_X})^{-1/2}.$$

Таким образом,

$$f_{X_1,X_2}(x_1,\ x_2)=(2\pi)^{-1/2}(\det \mathbf{R_X})^{-1/2}\exp\left(-\frac{1}{2}(x_1,\ x_2)\mathbf{R_X}^{-1}(x_1,\ x_2)'\right).$$

Найдем условную плотность $f_{X_2|X_1}(x_2|x_1)$. Положим $\mathbf{R_X}=\|r_{ij}\|$, $Q=\|q_{ij}\|$. По доказанному

$$\mathbf{R_X}=Q^{-1},\quad \det \mathbf{R_X}=(\det Q)^{-1},\quad \det Q=q_{11}q_{22}-q_{12}^2.$$

$$f_{X_1,X_2}(x_1,\ x_2)=(2\pi)^{-1}(\det Q)^{1/2}\exp\left(-\frac{1}{2}\ (q_{11}x_1^2+2q_{12}x_1x_2+q_{22}x_2^2)\right).$$

Так как сл. в. X_1 нормальна, $\mathbf{M}X_1=0$, $\mathbf{D}X_1=r_{11}=q_{22}/\det Q$, то

$$f_{X_1}(x_1)=(2\pi q_{22}^{-1}\det Q)^{1/2}\exp\left(-\frac{1}{2}(\det Q/q_{22})x_1^2\right),$$

$$f_{X_2|X_1}(x_2|x_1)=f_{X_1,X_2}(x_1,\ x_2)/f_{X_1}(x_1)=(2\pi q_{22}^{-1})^{-1/2}\times$$

$$\times\exp\left(-\frac{1}{2}(q_{11}x_1^2+2q_{12}x_1x_2+q_{22}x_2^2-q_{11}x_1^2+(q_{12}^2/q_{22})x_1^2)\right)=$$

$$=(2\pi q_{22}^{-1})^{-1/2}\exp\left(-\frac{1}{2}((q_{12}^2/q_{22})x_1^2+2q_{12}x_1x_2+q_{22}x_2^2)\right)=$$

$$=(2\pi q_{22}^{-1})^{-1/2}\exp\left(-\frac{1}{2}((x_2-(-q_{12}/q_{22})x_1)^2/q_{22}^{-1})\right),$$

что представляет собой нормальную плотность со средним $-(q_{12}/q_{22})x_1$ и дисперсией q_{22}^{-1}. Отсюда

$$\mathbf{M}(X_2|X_1)=-(q_{12}/q_{22})X_1.$$

Проведенные рассуждения дословно переносятся на пространство любой размерности: *n-мерное невырожденное нормальное распределение* с нулевым средним задается плотностью вида

$$f(x_1,\ \ldots,\ x_n)=(2\pi)^{-n/2}(\det Q)^{1/2}\exp\left(-\frac{1}{2}\ \mathbf{x}Q\mathbf{x}'\right),\quad \mathbf{x}=(x_1,\ \ldots,\ x_n),$$

где Q — произвольная симметричная положительно определенная матрица. Нормальный случайный вектор $\mathbf{X}=(X_1,\ \ldots,\ X_n)$ с плотностью $f(x_1,\ldots,x_n)$ имеет нулевое математическое ожидание и матрицу ковариаций $\mathbf{R_x}=Q^{-1}$. Если C — ортогональная матрица, приводящая $\mathbf{R_x}$ к диагональной форме: $C'\mathbf{R_x}\,C=\Lambda$, Λ — диагональная матрица с положительными числами λ_i, $i=1,\ \ldots,\ n$, на главной диагонали, то вектор $\mathbf{Y}=\mathbf{X}C$ состоит из независимых нормальных компонент. Любая нетривиальная линейная комбинация компонент вектора $\mathbf{X}$ имеет нормальное распределение. Условная плотность $f_{X_n|X_1,\ldots,X_{n-1}}(x_n|x_1, \ldots, x_{n-1})$ нормальна со средним

$$-(q_{1n}/q_{nn})x_1 - \ldots - (q_{n-1,n}/q_{nn})x_n$$

и дисперсией q_{nn}^{-1}, так что

$$\mathbf{M}(X_n|X_1, \ldots, X_{n-1}) = -(q_{1n}/q_{nn})X_1 - \ldots - (q_{n-1,n}/q_{nn})X_n.$$

Сравнивая последнее соотношение с 8.36, получаем важный вывод: для нормального случайного вектора наилучшая в среднеквадратичном нелинейная оценка X_n по $X_1, \ldots, X_{n-1}$ совпадает с линейной.

9.19. Пример (наилучшее несмещенное оценивание). Одна из постановок задач математической статистики состоит в следующем. Дана последовательность наблюдений $\mathbf{x}=(x_1, \ldots, x_n)$, которые рассматриваются как значения, принятые последовательностью $\mathbf{X}=(X_1, \ldots, X_n)$ независимых одинаково распределенных случайных величин, распределение вероятностей которых определено с точностью до неизвестного параметра θ. Требуется построить оценку $\hat{\theta}$ параметра по результатам наблюдений $x_1, \ldots$ $\ldots, x_n$, т. е. такую функцию $\hat{\theta}=S(x_1, \ldots, x_n)$, для которой в каком-то смысле можно было говорить о приближенном равенстве $\hat{\theta} \approx \theta$. Типичный пример такого рода — оценка неизвестной вероятности успеха в n испытаниях Бернулли. В более общей постановке речь идет об оценке некоторой функции $\psi(\theta)$ неизвестного параметра.

Пусть для оценки параметрической функции $\psi(\theta)$ предложена некоторая *статистика* — функция результатов наблюдений $S(x_1, \ldots, x_n)$. Как характеризовать погрешность оценивания $S(x_1, \ldots, x_n) - \psi(\theta)$, имея в виду, что в качестве $x_1, \ldots, x_n$ могут выступать любые значения, принятые в результате проведения опыта случайными величинами $X_1, \ldots, X_n$. Фактически речь идет об отклонении значения сл. в. $S(X_1, \ldots, X_n)$ от некоторой (неизвестной) постоянной $\psi(\theta)$. Мы уже знакомы с одной такой мерой разброса — это среднеквадратичное расстояние

$$(\mathbf{M}(S(X_1, \ldots, X_n) - \psi(\theta))^2)^{1/2}.$$

Если оценка $S(X_1, \ldots, X_n)$ не имеет смещения ($\mathbf{M}S(X_1, \ldots, X_n) = \psi(\theta)$), то среднеквадратическое отклонение превращается в дисперсию $\mathbf{D}S(X_1, \ldots, X_n)$ и можно ставить вопрос о построении *несмещенной оценки* с минимальной дисперсией.

В математической статистике обычно не обращаются к исходному вероятностному пространству, на котором определены сл. в. $X_1, \ldots, X_n$, а имеют дело с пространством $(R^n, \mathscr{B}^n, \mathbf{P}_\theta)$, порожденным сл. в. $X_1, \ldots, X_n$; здесь $\mathscr{B}^n$ — класс подмножеств $B \subset R^n$, таких, что множество $\{\omega : (X_1(\omega), \ldots, X_n(\omega)) \in B\}$ является событием (см. определение $\mathscr{B}^n$ в 6.19, подробнее об этом будет идти речь в § 10), $\mathbf{P}_\theta$ — вероятностная мера на $\mathscr{B}^n$, представляющая собой распределение вероятностей случайного вектора $\mathbf{X}=(X_1, \ldots, X_n)$. Поскольку истинное значение параметра θ не известно, то мы стоим перед необходимостью рассматривать все допустимые значения θ в качестве возможных, а это фактически приводит к целому семейству вероятностных пространств $\{(R^n, \mathscr{B}^n, \mathbf{P}_\theta)\}$, зависящему от параметра θ. В этом семействе переменный только один элемент — распределение вероятностей $\mathbf{P}_\theta$. Это подводит к определению *статистической модели* как объекта $((R^n, \mathscr{B}^n), \{\mathbf{P}_\theta\})$, состоящего из *выборочного пространства* $(R^n, \mathscr{B}^n)$ и семейства вероятностных мер $\{\mathbf{P}_\theta\}$.

Функции результатов наблюдений $S(x_1, \ldots, x_n)$ называются *статистиками* (предполагается, что для любого промежутка $I_{a,b}$ $\{(x_1, \ldots, x_n) : S(x_1, \ldots, x_n) \in I_{a,b}\} \in \mathscr{B}^n$), они становятся случайными величинами, когда фиксируется мера $\mathbf{P}_\theta$. В обозначениях мы отражаем это обстоятельство, переходя от $S(x_1, \ldots, x_n)$ к $S(X_1, \ldots, X_n)$ и употребляя символы $\mathbf{M}_\theta$, $\mathbf{D}_\theta$ для математического ожидания и дисперсии случайных величин с распределением $\mathbf{P}_\theta$.

Важно иметь в виду, что поскольку истинное значение параметра θ не известно, то качество любой оценки $S(x_1, \ldots, x_n)$ неизвестного $\psi(\theta)$ следует рассматривать во всей области допустимых значений θ. В некоторых случаях заранее ограничивают себя классом оценок, для которых некоторое свойство выполняется при всех θ. Таков, например, класс *несмещенных оценок*, когда выполняются равенства

$$\mathbf{M}_\theta S(X_1, \ldots, X_n) = \psi(\theta),$$

где θ пробегает все параметрическое множество. Задача *наилучшего несмещенного оценивания* заключается в нахождении такой функции $S(x_1, \ldots, x_n)$, удовлетворяющей уравнению несмещенности, для которой дисперсия $\mathbf{D}_\theta S(X_1, \ldots, X_n)$ была бы минимальна при каждом θ.

Ограничимся далее случаем дискретного распределения вероятностей $\mathbf{P}_\theta$. Пусть $p^\theta_{\mathbf{X}}(\mathbf{x}) = p^\theta_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ обозначает совместное распределение сл. в. $X_1, \ldots, X_n$, рассматриваемых относительно вероятностной модели $(R^n, \mathscr{B}^n, \mathbf{P}_\theta)$. Последующие построения опираются на предположение, что существует такая функция $T(x_1, \ldots, x_n)$, что условное распределение

$$p(\mathbf{x} \mid t) = p_{\mathbf{X}|T}(\mathbf{x} \mid t) = p^\theta_{\mathbf{X}}(\mathbf{x}) / p^\theta_T(t), \quad t = T(\mathbf{x}), \quad \mathbf{x} = (x_1, \ldots, x_n),$$

сл. в. $\mathbf{X}=(X_1, \ldots, X_n)$ при условии сл. в. $T=T(X_1, \ldots, X_n)$ одно и то же при всех значениях θ. С примерами такого рода мы встречались в задачах 5.1—5.5. Там же объяснялось, что статистика $T(\mathbf{x})$, обладающая указанным свойством, несет всю полезную информацию о неизвестном θ, содержащуюся в наблюдениях

$\mathbf{x}=(x_1, \ldots, x_n)$. По этой причине $T(\mathbf{x})$ называется *достаточной статистикой*.

Следующий подход, предложенный независимо Колмогоровым, Рао и Блекуэллом, позволяет по любой несмещенной оценке $S(\mathbf{x})$ построить несмещенную оценку $g(T(\mathbf{x}))$ — некоторую функцию достаточной статистики $T(\mathbf{x})$, для которой

$$\mathbf{D}_\theta g(T(\mathbf{X})) \leqslant \mathbf{D}_\theta S(\mathbf{X})$$

при всех θ, причем неравенство строгое, если только $g(T(\mathbf{x}))$ не совпадает с $S(\mathbf{X})$ с вероятностью 1 (по мере $\mathbf{P}_\theta$). Эта *улучшенная* несмещенная оценка, построенная по $S(\mathbf{X})$, имеет вид

$$g(T(\mathbf{X}))=\mathbf{M}_\theta(S(\mathbf{X}) \mid T(\mathbf{X})).$$

Прежде всего заметим, что

$$g(t)=\mathbf{M}_\theta(S(\mathbf{X}) \mid \{\omega : T(\mathbf{X}(\omega))=t\})=\sum_{\{\mathbf{x}\}} S(\mathbf{x})\, p(\mathbf{x} \mid t)$$

в самом деле не зависит от θ и, следовательно, $g(T(\mathbf{x}))$ действительно является функцией от наблюдений $\mathbf{x}=(x_1, \ldots, x_n)$. Далее,

$$\mathbf{M}_\theta g(T(\mathbf{X}))=\mathbf{M}_\theta \mathbf{M}_\theta(S(\mathbf{X}) \mid T(\mathbf{X}))=\mathbf{M}_\theta S(\mathbf{X})=\psi(\theta),$$

т. е. оценка $g(T(\mathbf{X}))$ несмещенная.

Наконец, неравенство для дисперсий

$$\| g(T(\mathbf{X}))-\psi(\theta) \|_\theta^2 \leqslant \| S(\mathbf{X})-\psi(\theta) \|_\theta^2$$

объясняется тем, что $g(T(\mathbf{X}))-\psi(\theta)$ — проекция (в гильбертовом пространстве случайных величин, рассматриваемых по мере $\mathbf{P}_\theta$) вектора $S(\mathbf{X})-\psi(\theta)$ на подпространство случайных величин, являющихся действительными функциями от $T(\mathbf{X})=T$. Докажем это неравенство, опираясь на свойства условных математических ожиданий:

$$\mathbf{D}_\theta S=\mathbf{M}_\theta(S-\psi(\theta))^2=\mathbf{M}_\theta(S-g(T)+g(T)-\psi(\theta))^2=$$

$$=\mathbf{M}_\theta(S-g(T))^2+2\mathbf{M}_\theta(S-g(T))(g(T)-\psi(\theta))+\mathbf{M}(g(T)-\psi(\theta))^2,$$

$$\mathbf{M}_\theta(S-g(T))(g(T)-\psi(\theta))=\mathbf{M}_\theta \mathbf{M}_\theta((S-g(T))(g(T)-\psi(\theta)) \mid T)=$$

$$=\mathbf{M}_\theta((g(T)-\psi(\theta))\,\mathbf{M}_\theta(S-g(T) \mid T))=\mathbf{M}_\theta((g(T)-\psi(\theta)) \times$$

$$\times(\mathbf{M}_\theta(S \mid T)-g(T)))=\mathbf{M}_\theta(g(T)-\psi(\theta))(g(T)-g(T))=0.$$

Таким образом,

$$\mathbf{D}_\theta S=\mathbf{M}_\theta(S-g(T))^2+\mathbf{D}_\theta g(T),$$

что и устанавливает требуемый результат.

9.20. Задача. Дана последовательность n испытаний Бернулли с неизвестной вероятностью успеха θ. Используя достаточную статистику $T(\mathbf{x})=x_1+\ldots+x_n$ (см. 5.1), построить улучшенную несмещенную оценку θ, исходя из несмещенной оценки $S(\mathbf{x})=x_1$ (которая сама как оценка θ никуда не годится, поскольку принимает лишь значения 0 и 1).

[Так как $\mathbf{M}_\theta X_1=\theta$, то $S(\mathbf{X})=X_1$ — действительно несмещенная оценка параметра θ. Далее имеем

$$g(t)=\mathbf{M}_\theta(X_1|\{\omega: T(\mathbf{X}(\omega))=t\})=\mathbf{P}_\theta(\{\omega: X_1(\omega)=1\}|\{\omega: T(\mathbf{X}(\omega))=t\})=$$

$$=p^\theta_{X_1,T}(1,\ t)/p^\theta_T(t)=\theta\cdot C_{n-1}^{t-1}\theta^{t-1}(1-\theta)^{n-t}/C_n^t\theta^t(1-\theta)^{n-t}=t/n,$$

$$g(T(\mathbf{x}))=T(\mathbf{x})/n=(x_1+\ldots+x_n)/n,$$

т. е. улучшенная оценка является частотой успехов.]

9.21. Задача. Дана последовательность $X_1, \ldots, X_n$ независимых пуассоновских случайных величин с одинаковым, но неизвестным параметром λ. Пользуясь достаточной статистикой $T(\mathbf{x})=x_1+\ldots+x_n$ (см. 5.2) и несмещенной оценкой $S(\mathbf{x})=x_1$, $\mathbf{x}=(x_1, \ldots, x_n)$, построить улучшенную оценку параметра λ.

$$\Big[\ g(t)=\mathbf{M}_\theta(X_1|\{\omega: T(\mathbf{X}(\omega))=t\})=\sum_{\{\mathbf{x}\}}x_1 p_{\mathbf{X}|T}(\mathbf{x}|t)=$$

$$=\sum_{\{\mathbf{x}\}}x_1(t!/(x_1!\ \ldots\ x_n!))n^{-t}=$$

$$=tn^{-1}\sum_{\{\mathbf{x}:\, x_1>0\}}\frac{(t-1)!}{(x_1-1)!\,x_2!\ \ldots\ x_n!}(n-1)^{-(t-1)}=tn^{-1},\ g(T)=Tn^{-1},$$

где суммирование ограничено условием $x_1+\ldots+x_n=t$ и где учтено, что $(t!/(x_1!\ldots x_n!))n^{-t}$ представляют собой полиномиальные вероятности, в сумме равные 1.]

9.22. Задача. Сл. в. $X_1, \ldots, X_n$ независимы и одинаково равномерно распределены с множеством значений $1, 2, \ldots, N$, где N — неизвестно. Найти какую-либо несмещенную оценку параметра N, зависящую от первого наблюдения. Построить улучшенную оценку, используя достаточную статистику $T(\mathbf{x})=\max(x_1, \ldots, x_n)$ (см. 5.3).

[Так как

$$\mathbf{M}_\lambda X_1=N^{-1}(1+2+\ldots+N)=N^{-1}(N(N+1)/2)=(N+1)/2,$$

то несмещенной будет оценка $S(\mathbf{x})=2x_1-1$. Используя 5.3, получаем

$$p^{(N)}_{X_1,T}(k,\ t)=\begin{cases}\mathbf{P}(\omega: X_1(\omega)=k)\,\mathbf{P}(\omega:\max(X_2(\omega), \ \ldots,\ X_k(\omega)),=t), k<t,\\ \mathbf{P}(\omega: X_1(\omega)=t)\,\mathbf{P}(\omega: X_2(\omega)\leqslant t,\ \ldots,\ X_n(\omega)\leqslant t),\ k=t,\end{cases}$$

$$p^{(N)}_{X_1|T}(k|t)=\begin{cases}t^{-1}(1-(1-t^{-1})^{n-1})(1-(1-t^{-1})^n)^{-1}, & k<t,\\ t^{-1}(1-(1-t^{-1})^n)^{-1}, & k=t.\end{cases}$$

Отсюда находим

$$\mathbf{M}_N(X_1|\{\omega : T(\mathbf{X}(\omega))=t\})=\sum_{k=1}^{t} kp^{(N)}_{X_1|T}(k|t)=$$

$$=(1/2)(t-1)(1-(1-t^{-1})^{n-1})+1)(1-(1-t^{-1})^n)^{-1},$$

$$g(t)=\mathbf{M}_N(2X_1-1|\{\omega : T(\mathbf{X}(\omega))=t\})=$$

$$=(t-(t-1)(1-t^{-1})^n)(1-(1-t^{-1})^n)^{-1},$$

или

$$g(t)=t(1-(1-t^{-1})^{n+1})/(1-(1-t^{-1})^n).$$

При больших t имеем приближенно $g(t)\approx t(n+1)/n$, т. е. наилучшая оценка есть немного подправленная статистика $T(\mathbf{x})$.]

9.23. Пример (мартингалы). Пусть $X_1, X_2, \ldots$ — некоторая последовательность дискретных сл. в. и $\mathbf{M}X_i=0$, $i=1, 2, \ldots$. Если $X_1, X_2, \ldots$ независимы, то при любом n имеет место равенство

$$\mathbf{M}(X_{n+1}|X_1, \ldots, X_n)=0, \quad n=1, 2, \ldots .$$

Оказывается, что и когда сл. в. $X_1, X_2, \ldots$ зависимы, но выполняются указанные равенства, последовательность $X_1, X_2, \ldots$ сохраняет некоторые общие черты с последовательностью независимых величин. Это позволяет перенести на них некоторые результаты, справедливые для независимых величин. Например, из этих соотношений вытекает, что сл. в. X_i, $i=1, 2, \ldots$, некоррелированы: при $m<n$ имеем (см. 9.8, 9.11)

$$\mathbf{M}(X_mX_n)=\mathbf{M}(\mathbf{M}(X_mX_n|X_1, \ldots, X_{n-1}))=$$

$$=\mathbf{M}(X_m\mathbf{M}(X_n|X_1, \ldots, X_{n-1}))=0.$$

Образуем последовательность нарастающих сумм, составленных из сл. в. $X_1, X_2, \ldots$: $Y_n=X_1+\ldots+X_n$, $n=1, 2, \ldots$. Для последовательности $Y_1, Y_2, \ldots$ имеем при любом n

$$\mathbf{M}(Y_{n+1}|X_1, \ldots, X_n)=\mathbf{M}(X_{n+1}|X_1, \ldots, X_n)+$$

$$+\mathbf{M}(Y_n|X_1, \ldots, X_n)=Y_n$$

и

$$\mathbf{M}(Y_{n+1}|Y_1, \ldots, Y_n)=Y_n, \quad n=1, 2, \ldots .$$

Последовательность сл. в. $Y_1, Y_2, \ldots$, удовлетворяющая этой системе равенств, называется *мартингальной*, или *мартингалом*. Заметим, что если $Y_1, Y_2, \ldots$ — мартингал, то последовательность сл. в. $X_n=Y_n-Y_{n-1}$, $n=1, 2, \ldots$, удовлетворяет равенствам

$$\mathbf{M}(X_{n+1}|X_1, \ldots, X_n)=0, \quad n=1, 2, \ldots .$$

Мартингальная теория получила в настоящее время значительное развитие, ее методы оказались чрезвычайно плодотворными при исследовании многих вопросов теории вероятностей. Первоначально мартингалы появились при изучении роли систем в азартных играх. Уже простейшая игра в орлянку предоставляет возможности для выбора различных стратегий. Например, можно менять ставку. Опишем эту игру последовательностью $X_1, X_2, \ldots$ независимых сл. в., принимающих значения ± 1 с равными вероятностями. Естественно предположить, что выбор ставки делается на основании предшествующего хода игры и потому капитал игрока (выигрывающего при выпадении 1) после $(n+1)$-й партии можно представить в виде

$$S_{n+1} = S_n + X_{n+1} b_n (X_1, \ldots, X_n),$$

где b_n — сумма, разыгрываемая в $(n+1)$-й партии. Очевидно,

$$\mathbf{M}(S_{n+1} | X_1, \ldots, X_n) = \mathbf{M}(S_n + X_{n+1} b_n (X_1, \ldots, X_n) | X_1, \ldots, X_n) =$$
$$= \mathbf{M}(S_n | X_1, \ldots, X_n) + b_n (X_1, \ldots, X_n) \mathbf{M}(X_{n+1} | X_1, \ldots, X_n) = S_n,$$
$$\mathbf{M}(S_{n+1} | S_1, \ldots, S_n) = \mathbf{M}(\mathbf{M}(S_{n+1} | X_1, \ldots, X_n) | S_1, \ldots, S_n) = S_n,$$

т. е. последовательность $S_1, S_2, \ldots$ значений капитала — мартингал. Таким образом, вне зависимости от системы выбора ставки капитал игрока в следующей партии в среднем остается на уровне его значения в предшествующей партии.

9.24. Пример (ветвящийся процесс). Представим себе следующий процесс размножения частиц. Пусть в начальный момент времени имеется одна частица, которая к моменту 1 дает случайное число Z_1 таких же частиц, а сама выбывает из дальнейшего процесса размножения. К моменту 2 каждая из Z_1 частиц первого поколения, размножаясь независимо от остальных, дает случайное число потомков с тем же законом распределения, что и Z_1. Обозначим Z_2 суммарное число потомков второго поколения. Продолжая указанный процесс, получаем последовательность сл. в. $Z_0=1, Z_1, Z_2, \ldots$, которая называется ***ветвящимся процессом Гальтона — Ватсона***.

Формальное построение процесса Гальтона — Ватсона можно провести следующим образом. Пусть $X_i^{(j)}$, $i, j=1, 2, \ldots$, — независимые в совокупности одинаково распределенные случайные величины с целыми неотрицательными значениями (существование такого семейства сл. в. вытекает из построений § 6, если расположить сл. в. $X_i^{(j)}$ в последовательность, нумеруя их, например, в порядке возрастания $i+j$, а при $i+j=\mathrm{const}$ в произвольном порядке). Положим

$$Z_{n+1} = X_1^{(n+1)} + \ldots + X_{Z_n}^{(n+1)}, \quad n=0, 1, 2, \ldots, \quad Z_0 = 1.$$

Последовательность Z_n, $n=0, 1, \ldots$, образует, очевидно, марковскую цепь:

$$p_{Z_{n+1} | Z_0, \ldots, Z_n} (i_{n+1} | i_0, \ldots, i_n) = \mathbf{P}(\omega : X_1^{(n+1)}(\omega) + \ldots + X_{i_n}^{(n+1)}(\omega) = i_{n+1}).$$

Отсюда имеем (ср. 9.6)

$$M(Z_{n+1}|Z_0, Z_1, \ldots, Z_n) = Z_n \cdot \mu, \quad \mu = MZ_1.$$

Следовательно, последовательность $W_n = Z_n/\mu^n$, $n=0, 1, 2, \ldots$, — мартингал.

§ 10*).

ИЗМЕРИМЫЕ ФУНКЦИИ И ИНТЕГРАЛ

Пусть Ω — некоторое множество, $\mathscr{A}$ — некоторая σ-алгебра его подмножеств. Пара $(\Omega, \mathscr{A})$ называется *измеримым пространством*. Действительная функция $X(\omega)$, рассматриваемая на измеримом пространстве $(\Omega, \mathscr{A})$, называется *измеримой по Борелю*, если $X^{-1}(I) \equiv \{\omega : X(\omega) \in I\} \in \mathscr{A}$ для любого промежутка I. На самом деле для измеримой функции $X^{-1}(B) \in \mathscr{A}$ для любого борелевского множества B. Проверим это утверждение. Обозначим $\mathscr{B}'$ класс всех подмножеств $B' \subseteq R$, таких, что $X^{-1}(B') \in \mathscr{A}$, и покажем, что $\mathscr{B}' \supseteq \mathscr{B}$. Если $B' \in \mathscr{B}'$, т. е. $X^{-1}(B') \in \mathscr{A}$, то и $\overline{X^{-1}(B')} \in \mathscr{A}$. Но $X^{-1}(\bar{B}') = \overline{X^{-1}(B')}$, так что $\bar{B}' \in \mathscr{B}'$. Пусть $B_i' \in \mathscr{B}'$, т. е. $X^{-1}(B_i') \in \mathscr{A}$, $i=1, 2, \ldots$. Тогда $\bigcup_{(i)} X^{-1}(B_i') \in \mathscr{A}$ и, поскольку

$$\bigcup_{(i)} X^{-1}(B_i') = X^{-1}\left(\bigcup_{(i)} B_i'\right),$$

получаем $\bigcup_{(i)} B_i' \in \mathscr{B}'$. Таким образом, класс $\mathscr{B}'$ замкнут относительно перехода к дополнению и счетному объединению, к тому же $\varnothing$, $R \in \mathscr{B}'$. Следовательно, $\mathscr{B}'$ — σ-алгебра. Поскольку $\mathscr{B}'$ содержит промежутки, то $\mathscr{B}'$ включает и порождаемую промежутками σ-алгебру борелевских множеств $\mathscr{B}$. Из проведенных рассуждений видно, что для измеримости функции $X(\omega)$ достаточно, чтобы были измеримы прообразы какого-либо класса подмножеств, порождающего $\mathscr{B}$. Например, достаточно, чтобы $X^{-1}((-\infty, x]) \in \mathscr{A}$ при любом x.

Удобно пополнить числовую прямую бесконечно удаленными точками: $\bar{R} = \{R, \infty, -\infty\}$, расширить систему борелевских множеств, добавляя к $\mathscr{B}$ подмножества $B \cup \{\infty\}$, $B \cup \{-\infty\}$, $B \cup \cup \{\infty, -\infty\}$ при всех $B \in \mathscr{B}$, и обобщить понятие измеримой функции, допуская в качестве значений $\pm\infty$ и предполагая $X^{-1}(\infty)$, $X^{-1}(-\infty) \in \mathscr{A}$. Введение бесконечно удаленных точек облегчает обращение с предельными операциями: монотонная последовательность всегда имеет предел, хотя он уже может быть бесконечным, всякое числовое множество имеет точные верхнюю

*) См. сноску на с. 101.

и нижнюю грани, которые, в частности, могут равняться $\pm\infty$. Бесконечно удаленные точки включаются в порядковую структуру прямой ($-\infty<x<\infty$ для любого $x\in R$) и в арифметическую структуру в предположении, что избегаются бессмысленные выражения типа $\infty-\infty$, ∞/∞ и др. Вместе с тем выражения $0\cdot\infty$, $\infty\cdot 0$ полезно отнести к осмысленным, приняв $0\cdot\infty=\infty\cdot 0=0$.

Класс измеримых функций замкнут относительно предельного перехода. В самом деле, пусть $X_n(\omega)$, $n=1, 2, \ldots$, измеримы и $Y(\omega)=\sup_n X_n(\omega)$. Так как при любом t $\{\omega : X_n(\omega)\in(-\infty, t]\}\in\mathscr{A}$, то

$$\{\omega : Y(\omega)\in(-\infty, t]\}=\bigcap_{n=1}^{\infty}\{\omega : X_n(\omega)\in(-\infty, t]\}\in\mathscr{A}.$$

Точно так же

$$Y^{-1}(\infty)=\bigcap_{N=1}^{\infty}\bigcup_{n=1}^{\infty}\{\omega : X_n(\omega)>N\}\in\mathscr{A},\quad Y^{-1}(-\infty)\in\mathscr{A}.$$

Следовательно, функция $Y(\omega)$ измерима. Аналогично устанавливается, что измерим инфинум (можно перейти к супремуму измеримых вместе с $X_n(\omega)$ функций $-X_n(\omega)$). В таком случае измеримы верхний предел последовательности измеримых функций

$$\overline{\lim_{n\to\infty}}\, X_n(\omega)=\inf_n \sup_{k\geqslant n} X_k(\omega)$$

и нижний предел. Если существует $\lim\limits_{n\to\infty} X_n(\omega)=X(\omega)$, то $X(\omega)$ также измерима.

Измеримая функция с конечным числом значений

$$h(\omega)=\sum_{i=1}^{n} a_i I_{A_i}(\omega),$$

где $a_i\in\overline{R}$, $A_i\in\mathscr{A}$, $i=1, \ldots, n$, $A_iA_j=\varnothing$ при $i\neq j$ называется *простой*. Если число значений измеримой функции счетно (в написанной выше формуле надо положить $n=\infty$), то функция называется *элементарной*. Любую измеримую функцию $X(\omega)$ можно представить как предел последовательности простых функций:

$$X^{(n)}(\omega)=\sum_{k=-n^2}^{n^2} kn^{-1} I_{\{\omega : kn^{-1}<X(\omega)\leqslant(k+1)n^{-1}\}}(\omega)+$$
$$+nI_{\{\omega : X(\omega)>n\}}(\omega)-nI_{\{\omega : X(\omega)\leqslant -n\}}(\omega).$$

Для дальнейшего заметим, что в случае неотрицательной $X(\omega)$ последовательность

$$X^{(n)}(\omega)=\sum_{\{0\leqslant k\leqslant 2^{2n}\}} k2^{-n} I_{\{\omega : k2^{-n}<X(\omega)\leqslant(k+1)2^{-n}\}}(\omega)$$

сходится к $X(\omega)$ не убывая.

Легко видеть, что класс простых функций замкнут относительно арифметических операций. Применяя предельный переход, получаем, что это утверждение верно для всего класса измеримых функций. Отсюда получаем также, что вместе с $X(\omega)$ измеримы функции

$$X^{+}(\omega)=X(\omega)\cdot I_{\{\omega\,:\,X(\omega)\geqslant 0\}}(\omega),\quad X^{-}=(-X)^{+},\quad |X|=X^{+}+X^{-}.$$

Предположим, что на измеримом пространстве $(\Omega, \mathscr{A})$ определена счетно-аддитивная мера μ. Когда $\mu(\Omega)=1$, тройка $(\Omega, \mathscr{A}, \mu)$ называется вероятностным пространством, любая измеримая функция — случайной величиной. Для случайных величин определено понятие математического ожидания. Это понятие в общей теории пространств с мерой называется *интегралом Лебега*. Существуют различные подходы к его определению. В § 7 был использован предельный переход от элементарных функций к равномерно приближаемым ими произвольным измеримым. Ниже используется, возможно не столь наглядный, но технически более совершенный подход.

Определим *интеграл Лебега* для неотрицательной простой функции формулой

$$\int h(\omega)\,\mu(d\omega)=\sum_{i=1}^{n} a_i\mu(A_i)\quad (0\cdot\infty=\infty\cdot 0=0),$$

а для произвольной простой положим

$$\int h(\omega)\,\mu(d\omega)=\int h^{+}(\omega)\,\mu(d\omega)-\int h^{-}(\omega)\,\mu(d\omega)$$

в предположении, что хотя бы один из интегралов справа конечен. Простую функцию всегда можно представить в виде линейной комбинации индикаторов с попарно различными коэффициентами: $a_i\neq a_j$ при $i\neq j$. Отсюда, принимая во внимание аддитивность меры μ, заключаем, что определение корректно — не зависит от представления простой функции как комбинации индикаторов. Интеграл от неотрицательной функции существует всегда, хотя и может равняться $+\infty$. Простая функция называется *интегрируемой*, если интеграл существует и конечен.

10.1. Задача. Проверить элементарные свойства интеграла от простых неотрицательных функций: аддитивность, однородность, монотонность $\left(\int h_1(\omega)\,\mu(d\omega)\leqslant\int h_2(\omega)\,\mu(d\omega)\text{ при } h_1(\omega)\leqslant h_2(\omega)\right)$. [Заметим, что простую функцию можно записать в виде

$$h(\omega)=\sum_{i=1}^{n} a_i I_{A_i}(\omega),$$

где измеримые множества A_i, $i=1, \ldots, n$, образуют разбиение Ω

(т. е. $\bigcup_{i=1}^{n} A_i=\Omega$, $A_iA_j=\varnothing$ при $i\neq j$), добавляя, если это необходимо, член вида $0\cdot I_A(\omega)$. Пусть

$$h_k(\omega)=\sum_{i=1}^{n_k} a_i^{(k)} I_{A_i^{(k)}}(\omega), \quad k=1,\ 2,$$

— простые функции, представленные в указанном виде. В таком случае имеем

$$1=\sum_{i=1}^{n_2} I_{A_i^{(2)}}(\omega), \quad I_{A_i^{(1)}}(\omega)=\sum_{j=1}^{n_2} I_{A_i^{(1)}} I_{A_j^{(2)}}=\sum_{j=1}^{n_2} I_{A_i^{(1)}A_j^{(2)}},$$

$$h_1(\omega)=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} a_i^{(1)} I_{A_i^{(1)}A_j^{(2)}}, \quad \mu(A_i^{(1)})=\sum_{j=1}^{n_2}\mu(A_i^{(1)}A_j^{(2)})$$

и аналогично

$$h_2(\omega)=\sum_{j=1}^{n_2}\sum_{i=1}^{n_1} a_j^{(2)} I_{A_i^{(1)}A_j^{(2)}}, \quad \mu(A_j^{(2)})=\sum_{i=1}^{n_1}\mu(A_i^{(1)}A_j^{(2)}),$$

откуда

$$\int h_1(\omega)\,\mu(d\omega)=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} a_i^{(1)}\mu(A_i^{(1)}A_j^{(2)}),$$

$$\int h_2(\omega)\,\mu(d\omega)=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} a_j^{(2)}\mu(A_i^{(1)}A_j^{(2)}).$$

Предположим, что $h_1(\omega)\leqslant h_2(\omega)$. Тогда для пар i, j, таких, что $A_i^{(1)}A_j^{(2)}\neq\varnothing$, имеем $a_i^{(1)}\leqslant a_j^{(2)}$. Сравнивая почленно суммы в правых частях, получаем

$$\int h_1(\omega)\,\mu(d\omega)\leqslant\int h_2(\omega)\,\mu(d\omega).$$

Складывая левые и правые части выписанных выше соотношений, находим

$$\int h_1(\omega)\,\mu(d\omega)+\int h_2(\omega)\mu(d\omega)=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(a_i^{(1)}+a_j^{(2)})\,\mu(A_i^{(1)}A_j^{(2)}),$$

что вместе с представлением

$$h_1(\omega)+h_2(\omega)=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(a_i^{(1)}+a_j^{(2)})\,I_{A_i^{(1)}A_j^{(2)}}$$

доказывает аддитивность интеграла. Однородность интеграла

$$\int ch(\omega)\mu(d\omega)=c\int h(\omega)\mu(d\omega)$$

очевидна.]

10.2. Задача. Перенести утверждения 10.1 на произвольные простые функции.

[В предположении, что каждый из интегралов

$$\int h_1(\omega)\mu(d\omega),\ \int h_2(\omega)\mu(d\omega)$$

существует, доказательство монотонности интеграла для неотрицательных простых функций переносится на простые функции, принимающие значения любого знака. Это же замечание относится к доказательству однородности и аддитивности в предположении, что суммы

$$\int h_1(\omega)\mu(d\omega)+\int h_2(\omega)\mu(d\omega),\ h_1(\omega)+h_2(\omega)$$

имеют смысл. Последнее предположение можно ослабить, если рассматривать функции, определенные всюду, кроме, быть может, некоторого множества меры нуль, и допускать неопределенность суммы $h_1(\omega)+h_2(\omega)$ на некотором множестве μ — меры нуль.]

10.3. Задача. Пусть $h(\omega)$ — неотрицательная простая функция. Положим для любого измеримого множества A

$$\int h(\omega)I_A(\omega)\mu(d\omega)=\int_A h(\omega)\mu(d\omega)\equiv\mathcal{J}_h(A),$$

где слева стоит интеграл от простой неотрицательной функции $h(\omega)I_A(\omega)$. Показать, что $\mathcal{J}_h(A)$ — счетно-аддитивная мера на $(\Omega, \mathcal{A})$.

[Для $h(\omega)=b\cdot I_B(\omega)$ получаем $\mathcal{J}_h(A)=b\cdot\mu(AB)$, и утверждение очевидно. Произвольная простая функция — сумма функций указанного вида, так что в общем случае $\mathcal{J}_h(A)$ представляется как сумма мер и, следовательно, является мерой.]

10.4. Задача. Для любой простой функции $h(\omega)$ и любого множества $A\in\mathcal{A}$ определим интеграл от $h(\omega)$ по множеству A как

$$\int_A h(\omega)\mu(d\omega)=\int h(\omega)I_A(\omega)\mu(d\omega)$$

при условии, что интеграл справа существует. Проверить элементарные свойства интеграла по множеству A от простых функций. Показать, что интеграл как функция множества счетно-аддитивен и непрерывен (предполагается, что интеграл от h существует).

$$\Big[\int_A (h_1(\omega)+h_2(\omega))\mu(d\omega)=\int (h_1(\omega)I_A(\omega)+h_2(\omega)I_A(\omega))\mu(d\omega)=$$

$$= \int h_1(\omega) I_{A_1}(\omega)\mu(d\omega) + \int h_2(\omega) I_{A_1}(\omega)\mu(d\omega) = \int_A h_1(\omega)\mu(d\omega) +$$

$$+ \int_A h_2(\omega)\mu(d\omega), \quad \int_A ch(\omega)\mu(d\omega) = c\int_A h(\omega)\mu(d\omega),$$

где предполагается, что интегралы

$$\int_A h_1(\omega)\mu(h\omega), \quad \int_A h_2(\omega)\mu(d\omega)$$

существуют, а их сумма определена ($\neq\infty, -\infty$).

Если $h_1(\omega) \leqslant h_2(\omega)$ на множестве A, то, применяя 10.2 к функциям $h_1(\omega) I_A(\omega)$, $h_2(\omega) I_A(\omega)$, получаем

$$\int_A h_1(\omega)\mu(d\omega) \leqslant \int_A h_2(\omega)\mu(d\omega).$$

Счетная аддитивность и непрерывность функции множеств

$$\mathcal{J}_h(A) = \int_A h(\omega)\mu(d\omega)$$

вытекает из представления

$$\mathcal{J}_h(A) = \mathcal{J}_{h^+}(A) - \mathcal{J}_{h^-}(A), \quad h(\omega) = h^+(\omega) - h^-(\omega)$$

и счетной аддитивности и непрерывности мер $\mathcal{J}_{h^+}(A)$, $\mathcal{J}_{h^-}(A)$.]

10.5. Задача. Пусть $h(\omega)$ — конечная неотрицательная простая функция. Вывести неравенство

$$\int h(\omega)\mu(d\omega) \leqslant \max_\omega \{h(\omega)\}\mu(\omega : h(\omega) > \varepsilon) + \varepsilon\mu(\omega : 0 < h(\omega) \leqslant \varepsilon).$$

$$\left[\ \int h(\omega)\mu(d\omega) = \int_\Omega h(\omega)\mu(d\omega) = \sum_{i=1}^{3}\int_{A_i} h(\omega)\mu(d\omega),\right.$$

$$A_1 = \{\omega : h(\omega) > \varepsilon\}, \quad A_2 = \{\omega : 0 < h(\omega) \leqslant \varepsilon\}, \quad A_3 = \{\omega : h(\omega) = 0\}.$$

На множестве A_1 воспользуемся неравенством $h(\omega) \leqslant \max_\omega \{h(\omega)\}$, на множестве A_2 — неравенством $h(\omega) \leqslant \varepsilon$, и, применяя свойство монотонности интеграла, получаем требуемое.]

10.6. Задача. Пусть $h_n(\omega)$, $n=1, 2, \ldots$, — невозрастающая последовательность простых функций, такая, что

$$\lim_{n\to\infty} h_n(\omega) = 0, \quad \int h_1(\omega)\mu(d\omega) < \infty.$$

Показать, что

$$\lim_{n\to\infty} \int h_n(\omega)\mu(d\omega) = 0.$$

[Так как $0 \leqslant h_n \leqslant h_1$, то h_n интегрируема и $\mu(\omega : h_n(\omega) = \infty) = 0$. Поскольку значение интеграла не меняется при изменении функции на множестве μ-меры нуль, можно без ущерба для общности предположить, что $h_n(\omega) < \infty$, $n=1, 2, \ldots$. Полагая $c = \max_\omega \{h_1(\omega)\}$ и учитывая, что $h_1(\omega) \geqslant h_2(\omega) \geqslant \ldots \geqslant 0$, из неравенства 10.5 получаем

$$0 \leqslant \int h_n(\omega)\mu(d\omega) \leqslant c\mu(\omega : h_n(\omega) > \varepsilon) + \varepsilon\mu(\omega : 0 < h_n(\omega) \leqslant \varepsilon) \leqslant$$
$$\leqslant c\mu(\omega : h_n(\omega) > \varepsilon) + \varepsilon\mu(\omega : h_1(\omega) > 0).$$

Заметим, что из конечности интеграла от $h_1(\omega)$ вытекает конечность $\mu(\omega : h_1(\omega) > 0)$, а из монотонной сходимости последовательности $h_n(\omega)$ к нулю следует, что

$$\{\omega : h_n(\omega) > \varepsilon\} \supseteq \{\omega : h_{n+1}(\omega) > \varepsilon\}, \quad \bigcap_{i=1}^{\infty} \{\omega : h_n(\omega) > \varepsilon\} = \varnothing,$$

причем $\mu(\omega : h_n(\omega) > \varepsilon) \leqslant \mu(\omega : h_1(\omega) > 0) < \infty$. Переходя в неравенстве к верхнему пределу по n и используя свойство непрерывности меры μ, получаем

$$0 \leqslant \overline{\lim_{n\to\infty}} \int h_n(\omega)\mu(d\omega) \leqslant \varepsilon\mu(\omega : h_1(\omega) > 0), \quad \varepsilon > 0].$$

10.7. Задача. Пусть $h_n(\omega)$, $n=1, 2, \ldots$, — невозрастающая последовательность неотрицательных простых функций сходится к простой функции $h(\omega)$ и $\int h_1(\omega)\mu(d\omega) < \infty$. Показать, что

$$\lim_{n\to\infty} \int h_n(\omega)\mu(d\omega) = \int h(\omega)\mu(d\omega) < \infty.$$

[Последовательность $h_n(\omega) - h(\omega)$, $n=1, 2, \ldots$, неотрицательных простых функций не возрастает, сходится к нулю и

$$0 \leqslant \int (h_1(\omega) - h(\omega))\mu(d\omega) = \int h_1(\omega)\mu(d\omega) - \int h(\omega)\mu(d\omega) < \infty.$$

Применяя 10.6, получаем требуемое.]

10.8. Задача. Пусть $h_n(\omega)$, $n=1, 2, \ldots$, — неубывающая последовательность неотрицательных простых функций сходится к простой функции $h(\omega)$. Показать, что

$$\lim_{n\to\infty} \int h_n(\omega)\mu(d\omega) = \int h(\omega)\mu(d\omega),$$

разобрав случаи конечного и бесконечного интеграла от $h(\omega)$. [Случай конечного интеграла от $h(\omega)$ разбирается так же, как в 10.7. Предположим, что

$$\int h(\omega)\mu(d\omega) = \sum_{i=1}^{n} a_i\mu(A_i) = \infty,$$

и пусть, например, $a_1\mu(A_1)=\infty$. Покажем, что

$$\int\limits_{A_1} h_n(\omega)\,\mu(d\omega) \to \infty, \quad n\to\infty.$$

Последовательность $\tilde{h}_n(\omega) = h_n(\omega)\,I_{A_1}(\omega)$, $n = 1, 2, \ldots$, не убывает и сходится к $a_1 I_{A_1}(\omega)$. Для любого $0<\alpha<a_1$ имеем

$$\{\omega : \tilde{h}_n(\omega)>\alpha\} \subseteq \{\omega : \tilde{h}_{n+1}(\omega)>\alpha\}, \quad \bigcup_{n=1}^{\infty} \{\omega : \tilde{h}_n(\omega)>\alpha\} = A_1,$$

и по свойству непрерывности меры

$$\mu(\omega : \tilde{h}_n(\omega)>\alpha) \to \mu(A_1), \quad n\to\infty.$$

Очевидно,

$$\int \tilde{h}_n(\omega)\,\mu(d\omega) \geqslant \alpha\mu(\omega : \tilde{h}_n(\omega)>\alpha).$$

Переходя в неравенстве к пределу по n, получаем

$$\lim_{n\to\infty} \int \tilde{h}_n(\omega)\,\mu(d\omega) \geqslant \alpha\mu(A_1).$$

Если $\mu(A_1)=\infty$, то все доказано, а если $\mu(A_1)<\infty$, то $a_1=\infty$, так что α можно выбрать как угодно большим.]

10.9. Задача. Пусть $h_n(\omega)$, $n=1, 2, \ldots$, — неубывающая (невозрастающая) последовательность простых функций, сходящаяся к простой функции $h(\omega)$, и $\int h_1(\omega)\,\mu(d\omega) > -\infty$ $\left(\int h_1\mu(d\omega)<\infty\right)$. Показать, что

$$\lim_{n\to\infty} \int h_n(\omega)\,\mu(d\omega) = \int h(\omega)\,\mu(d\omega) > -\infty \ (<\infty).$$

[Для неубывающей последовательности $h_n(\omega)$, $n=1, 2, \ldots$, последовательность $h_n^+(\omega)$, $n=1, 2, \ldots$, не убывает и сходится к $h^+(\omega)$, а последовательность $h_n^-(\omega)$, $n=1, 2, \ldots$, не возрастает и сходится к $h^-(\omega)$, причем $\int h_1^-(\omega)\,\mu(d\omega)<\infty$. Применяя 10.7, 10.8, получаем при $n\to\infty$

$$\int h_n^+(\omega)\,\mu(d\omega) \to \int h^+(\omega)\,\mu(d\omega), \quad \int h_n^-(\omega)\,\mu(d\omega) \to \int h^-(\omega)\,\mu(d\omega) < \infty,$$

и остается воспользоваться определением интеграла от простой функции.]

10.10. Задача. Пусть $h_n(\omega)$, $n=1, 2, \ldots$, — неубывающая последовательность простых функций, такая, что

$$\lim_{n\to\infty} h_n(\omega) \geqslant 0, \quad \int h_1(\omega)\,\mu(d\omega) > -\infty.$$

Показать, что

$$\lim_{n\to\infty} \int h_n^-(\omega)\,\mu(d\omega) = 0, \quad \lim_{n\to\infty} \int h_n(\omega)\,\mu(d\omega) \geqslant 0.$$

[Второе соотношение, очевидно, вытекает из первого. Из представления $h_n(\omega)=h_n^+(\omega)-h_n^-(\omega)$ следует, что последовательность $h_n^-(\omega)$, $n=1, 2, \ldots$, неотрицательных простых функций не возрастает и $\lim\limits_{n\to\infty} h_n^-(\omega)\leqslant 0$, а значит, $\lim\limits_{n\to\infty} h_n^-(\omega)=0$. Так как по условию $\int h_1^-(\omega)\,\mu(d\omega)<\infty$, то, применяя 10.6, получаем первое из требуемых соотношений.]

10.11. Задача. Пусть $h_n(\omega)$, $n=1, 2, \ldots$, — неубывающая последовательность простых функций, такая, что

$$\int h_1(\omega)\,\mu(d\omega)>-\infty, \quad \lim_{n\to\infty} h_n(\omega)\geqslant h(\omega)$$

где $h(\omega)$—простая функция, интеграл от которой существует. Показать, что

$$\lim_{n\to\infty}\int h_n(\omega)\,\mu(d\omega)\geqslant\int h(\omega)\,\mu(d\omega).$$

[Если $\int h(\omega)\,\mu(d\omega)<\infty$, то к последовательности $h_n(\omega)-h(\omega)$, $n=1, 2, \ldots$, применимо 10.10. Предположим, что $\int h(\omega)\,\mu(d\omega)=\infty$, и покажем, следуя схеме рассуждений 10.8, что в этом случае

$$\lim_{n\to\infty}\int h_n(\omega)\,\mu(d\omega)=\infty.$$

Пусть

$$\int h(\omega)\,\mu(d\omega)=\sum_{i=1}^{n} a_i\mu(A_i), \quad a_1\mu(A_1)=\infty, \quad \tilde{h}_n(\omega)=h_n(\omega)\,I_{A_1}(\omega).$$

Для любого $0<\alpha<a_1$ имеем

$$\int h_n(\omega)\,\mu(d\omega)\geqslant\alpha\mu(\omega: h_n(\omega)>\alpha)\geqslant\alpha\mu(\omega:\tilde{h}_n(\omega)>\alpha).$$

Последовательность $\tilde{h}_n(\omega)$, $n=1, 2, \ldots$, не убывает, $\lim\limits_{n\to\infty}\tilde{h}_n(\omega)\geqslant a_1$ при всех $\omega\in A_1$, $\tilde{h}_n(\omega)=0$ при $\omega\,\overline{\in}\,A_1$. Поэтому

$$\{\omega:\tilde{h}_n(\omega)>\alpha\}\subseteq\{\omega:\tilde{h}_{n+1}(\omega)>\alpha\}, \quad \bigcup_{n=1}^{\infty}\{\omega:\tilde{h}_n(\omega)>\alpha\}=A_1,$$

$$\lim_{n\to\infty}\mu(\omega:\tilde{h}_n(\omega)>\alpha)=\mu(A_1).$$

Таким образом,

$$\lim_{n\to\infty}\int h_n(\omega)\,\mu(d\omega)\geqslant\alpha\mu(A_1),$$

где либо $\mu(A_1)=\infty$, либо $0<\mu(A_1)<\infty$, но $a_1=\infty$, и следовательно, α можно взять сколько угодно большим.]

10.12. Задача. Пусть $h_n'(\omega)$, $h_n''(\omega)$, $n=1, 2, \ldots$, — две неубывающие последовательности простых неотрицательных функций, такие, что

$$\lim_{n\to\infty} h_n'(\omega) \geqslant \lim_{n\to\infty} h_n''(\omega).$$

Показать, что

$$\lim_{n\to\infty} \int h_n'(\omega)\,\mu(d\omega) \geqslant \lim_{n\to\infty} \int h_n''(\omega)\,\mu(d\omega).$$

[Так как при любом m $\lim\limits_{n\to\infty} h_n''(\omega) \geqslant h_m''(\omega)$, то

$$\lim_{n\to\infty} h_n'(\omega) \geqslant h_m''(\omega).$$

Применяя 10.11, получаем

$$\lim_{n\to\infty} \int h_n'(\omega)\,\mu(d\omega) \geqslant \int h_m''(\omega)\,\mu(d\omega).$$

Переходя в этом неравенстве к пределу по m, получаем требуемый результат.]

Интеграл Лебега от произвольной неотрицательной измеримой функции $X(\omega)$ определим как предел интегралов неубывающей последовательности $h_n(\omega)$, $n=1, 2, \ldots$, неотрицательных простых функций, сходящихся к $X(\omega)$:

$$\int X(\omega)\,\mu(d\omega) = \lim_{n\to\infty} \int h_n(\omega)\,\mu(d\omega).$$

Корректность определения вытекает из 10.12: если $h_n'(\omega)$, $h_n''(\omega)$, $n=1, 2, \ldots$, — две последовательности указанного вида, то из равенства

$$\lim_{n\to\infty} h_n'(\omega) = \lim_{n\to\infty} h_n''(\omega) = X(\omega)$$

следует равенство

$$\lim_{n\to\infty} \int h_n'(\omega)\,\mu(d\omega) = \lim_{n\to\infty} \int h_n''(\omega)\,\mu(d\omega).$$

Для произвольной измеримой функции $X(\omega)$ определим

$$\int X(\omega)\,\mu(d\omega) = \int X^+(\omega)\,\mu(d\omega) - \int X^-(\omega)\,\mu(d\omega)$$

в предположении, что хотя бы один из интегралов в правой части конечен. Функцию $X(\omega)$ назовем *интегрируемой*, если интеграл существует и конечен.

10.13. Задача. Показать, что для неотрицательных измеримых функций $X(\omega)$, $Y(\omega)$

$$\int X(\omega)\,\mu(d\omega)\geqslant 0,\quad \int cX(\omega)\,\mu(d\omega)=c\int X(\omega)\,\mu(d\omega),$$

$$\int (X(\omega)+Y(\omega))\,\mu(d\omega)=\int X(\omega)\,\mu(d\omega)+\int Y(\omega)\,\mu(d\omega),$$

а при условии $X(\omega)\geqslant Y(\omega)$

$$\int (X(\omega)-Y(\omega))\,\mu(d\omega)=\int X(\omega)\,\mu(d\omega)-\int Y(\omega)\,\mu(d\omega),$$

если хотя бы один из интегралов справа конечен, и

$$\int X(\omega)\,\mu(d\omega)\geqslant \int Y(\omega)\,\mu(d\omega).$$

[Первое соотношение очевидно, второе соотношение при $c>0$ и третье получаются предельным переходом от простых функций, соотношение

$$\int (-X(\omega))\,\mu(d\omega)=-\int X(\omega)\,\mu(d\omega)$$

представляет собой определение интеграла от отрицательной функции. При $X(\omega)\geqslant Y(\omega)$ по аддитивности получаем

$$\int X(\omega)\,\mu(d\omega)=\int (X(\omega)-Y(\omega)+Y(\omega))\,\mu(d\omega)=$$

$$=\int (X(\omega)-Y(\omega))\,\mu(d\omega)+\int Y(\omega)\,\mu(d\omega)\geqslant \int Y(\omega)\,\mu(d\omega),$$

что устанавливает два последних соотношения.

10.14. Задача. Пусть $0\leqslant X_1(\omega)\leqslant X_2(\omega)\leqslant\ldots$ — последовательность измеримых функций, $X(\omega)=\lim\limits_{n\to\infty} X_n(\omega)$. Построить неубывающую последовательность $h_n(\omega)$, $n=1, 2, \ldots$, неотрицательных простых функций, такую, что

$$h_n(\omega)\leqslant X_n(\omega),\quad n=1,\ 2,\ \ldots,\quad \lim_{n\to\infty} h_n(\omega)=X(\omega).$$

[Положим

$$g_n(\omega)=\sum_{k=0}^{n^2-1} kn^{-1} I_{\{\omega:\,kn^{-1}<X_n(\omega)\leqslant(k+1)n^{-1}\}}(\omega)+nI_{\{\omega:\,X_n(\omega)>n\}}(\omega).$$

Пусть $0\leqslant X(\omega)-X_n(\omega)<\varepsilon$ для всех $n>N=N(\varepsilon,\omega)$. При $X_n(\omega)\leqslant n$ имеем $0\leqslant X_n(\omega)-g_n(\omega)\leqslant n^{-1}$. Поэтому при n, ω, таких, что $n>\max(N(\omega), X_n(\omega))$, получаем $0\leqslant X(\omega)-g_n(\omega)<\varepsilon+n^{-1}$, что устанавливает сходимость последовательности $g_n(\omega)$, $n=1, 2, \ldots$, к $X(\omega)$ в точках ω, в которых $X(\omega)<\infty$. Аналогично устанавливается, что $g_n(\omega)\to\infty$, $n\to\infty$, если $X(\omega)=\infty$. Очевидно, что $g_n(\omega)\leqslant X_n(\omega)$, но не обязательно является монотонной последовательностью, зато $h_n(\omega)=\max\limits_{k\leqslant n}\{g_k(\omega)\}$ удовлетворяет всем поставленным требованиям, так как

$$g_n(\omega)\leqslant h_n(\omega)\leqslant \max_{k\leqslant n}\{X_k(\omega)\}=X_n(\omega).]$$

10.15. З а д а ч а (теорема о монотонной сходимости). Пусть $0\leqslant X_1(\omega)\leqslant X_2(\omega)\leqslant\ldots$ — последовательность измеримых функций, $X(\omega)=\lim\limits_{n\to\infty}X_n(\omega)$. Показать, что

$$\lim_{n\to\infty}\int X_n(\omega)\,\mu(d\omega)=\int X(\omega)\,\mu(d\omega).$$

[Воспользовавшись 10.14, построим неубывающую последовательность $h_n(\omega)$, $n=1, 2, \ldots$, неотрицательных простых функций, такую, что $h_n(\omega)\leqslant X_n(\omega)$, $\lim\limits_{n\to\infty}h_n(\omega)=X(\omega)$. Ввиду 10.13 и определения интеграла имеем

$$\int h_n(\omega)\,\mu(d\omega)\leqslant\int X_n(\omega)\,\mu(d\omega)\leqslant\int X(\omega)\,\mu(d\omega),$$

$$\lim_{n\to\infty}\int h_n(\omega)\,\mu(d\omega)=\int X(\omega)\,\mu(d\omega).]$$

10.16. З а д а ч а. Пусть $X(\omega)$ — неотрицательная измеримая функция. Положим для любого измеримого множества A

$$\mathcal{I}_X(A)=\int X(\omega)\,I_A(\omega)\,\mu(d\omega)\equiv\int_A X(\omega)\,\mu(d\omega).$$

Показать, что $\mathcal{I}_X(A)$ есть счетно-аддитивная мера на $(\Omega, \mathcal{A})$.

[Из 10.15 следует, что ряд из неотрицательных измеримых членов можно интегрировать почленно, так что при $A_iA_j=\varnothing$, $i\neq j$, получаем

$$\mathcal{I}_X\left(\bigcup_{i=1}^{\infty}A_i\right)=\int X(\omega)\sum_{i=1}^{\infty}I_{A_i}(\omega)\,\mu(d\omega)=$$

$$=\sum_{i=1}^{\infty}\int X(\omega)\,I_{A_i}(\omega)\,\mu(d\omega)=\sum_{i=1}^{\infty}\mathcal{I}_X(A_i).]$$

10.17. З а д а ч а. В предположении, что интеграл от $X(\omega)$ существует, показать, что функция множества

$$\mathcal{I}_X(A)=\int_A X(\omega)\,\mu(d\omega)\equiv\int X(\omega)\,I_A(\omega)\,\mu(d\omega)$$

счетно-аддитивна.

[Вытекает из 10.16 и представления $\mathcal{I}_X(A)=\mathcal{I}_{X^+}(A)-\mathcal{I}_{X^-}(A)$, являющегося определением интеграла от функции $X(\omega)I_A(\omega)$.]

10.18. З а д а ч а. Пусть $X(\omega)$, $Y(\omega)$ — неотрицательные измеримые функции. Показать, что если хотя бы один из интегралов справа конечен, то

$$\int (X(\omega)-Y(\omega))\,\mu(d\omega)=\int X(\omega)\,\mu(d\omega)-\int Y(\omega)\,\mu(d\omega).$$

[Положим $A=\{\omega : X(\omega)-Y(\omega)\geqslant 0\}$. Тогда

$$\int_A (X(\omega)-Y(\omega))\,\mu(d\omega)=\int (X(\omega)\,I_A(\omega)-Y(\omega)\,I_A(\omega))\,\mu(d\omega)=$$

$$=\int X(\omega)\,I_A(\omega)\,\mu(d\omega)-\int Y(\omega)\,I_A(\omega)\,\mu(d\omega),$$

где мы воспользовались 10.13, учитывая, что $X(\omega)I_A(\omega)\geqslant \geqslant Y(\omega)I_A(\omega)$. Аналогично

$$\int_{\bar A} (X(\omega)-Y(\omega))\,\mu(d\omega)=-\int_{\bar A} (Y(\omega)\,I_{\bar A}(\omega)-X(\omega)\,I_{\bar A}(\omega))\,\mu(d\omega)=$$

$$=-\int_{\bar A} Y(\omega)\,I_{\bar A}(\omega)\,\mu(d\omega)+\int_{\bar A} X(\omega)\,I_{\bar A}(\omega)\,\mu(d\omega).$$

Складывая полученные соотношения и пользуясь 10.17, приходим к требуемому результату.]

10.19. З а д а ч а. Проверить элементарные свойства интеграла: аддитивность, однородность, монотонность.

$$\Big[\int (X(\omega)+Y(\omega))\,\mu(d\omega)=\int ((X^+(\omega)+Y^+(\omega))-(X^-(\omega)+Y^-(\omega)))\,\mu(d\omega)=$$

$$=\int (X^+(\omega)+Y^+(\omega))\mu(d\omega)-\int (X^-(\omega)+Y^-(\omega))\,\mu(d\omega)=$$

$$=\int X^+(\omega)\,\mu(d\omega)+\int Y^+(\omega)\,\mu(d\omega)-\int X^-(\omega)\,\mu(d\omega)-\int Y^-(\omega)\,\mu(d\omega)=$$

$$=\int X(\omega)\,\mu(d\omega)+\int Y(\omega)\,\mu(d\omega),$$

где предполагается, что все выражения осмыслены и где мы воспользовались 10.18, 10.13 и определением интеграла.]

10.20. З а д а ч а. Пусть $X_1(\omega)\geqslant X_2(\omega)\geqslant\ldots\geqslant 0$ — измеримые функции, $X(\omega)=\lim\limits_{n\to\infty} X_n(\omega)$, $\int X_1(\omega)\,\mu(d\omega)<\infty$. Вывести из 10.15, что

$$\lim_{n\to\infty}\int X_n(\omega)\,\mu(d\omega)=\int X(\omega)\,\mu(d\omega).$$

[Последовательность $X_1(\omega)-X_n(\omega)$, $n=1, 2, \ldots$, не убывает и неотрицательна, так что

$$\lim_{n\to\infty}\int (X_1(\omega)-X_n(\omega))\,\mu(d\omega)=\int (X_1(\omega)-X(\omega))\,\mu(d\omega).$$

Применяя 10.13, имеем

$$\int (X_1(\omega)-X_n(\omega))\,\mu(d\omega)=\int X_1(\omega)\,\mu(d\omega)-\int X_n(\omega)\,\mu(d\omega),$$

$$\int (X_1(\omega)-X(\omega))\,\mu(d\omega)=\int X_1(\omega)\,\mu(d\omega)-\int X(\omega)\,\mu(d\omega),$$

откуда ввиду конечности интеграла от $X_1(\omega)$ получаем требуемый результат.]

10.21. Задача. Последовательность $X_n(\omega)$ измеримых функций не убывает (не возрастает) и

$$\int X_1(\omega)\,\mu(d\omega)>-\infty \quad \left(\int X_1(\omega)\,\mu(d\omega)<\infty\right).$$

Показать, что

$$\lim_{n\to\infty}\int X_n(\omega)\,\mu(d\omega)=\int \lim_{n\to\infty} X_n(\omega)\,\mu(d\omega).$$

Устанавливается аналогично 10.20, либо выводится из 10.20, 10.15 с помощью разложения $X_n(\omega)=X_n^+(\omega)-X_n^-(\omega)$.]

10.22. Задача (лемма Фату). Пусть $X_n(\omega)$, $n=1, 2, \ldots$, — неотрицательные измеримые функции. Показать, что

$$\int \underline{\lim}_{n\to\infty} X_n(\omega)\,\mu(d\omega)\leqslant \underline{\lim}_{n\to\infty}\int X_n(\omega)\,\mu(d\omega).$$

[Последовательность $Y_n(\omega)=\inf\limits_{k\geqslant n} X_k(\omega)$ измеримых неотрицательных функций не убывает и потому

$$\lim_{n\to\infty}\int Y_n(\omega)\,\mu(d\omega)=\int \lim_{n\to\infty} Y_n(\omega)\,\mu(d\omega).$$

Принимая во внимание, что

$$\int Y_n(\omega)\,\mu(d\omega)\leqslant\int X_n(\omega)\,\mu(d\omega),\quad \underline{\lim}_{n\to\infty} X_n(\omega)=\lim_{n\to\infty} Y_n(\omega),$$

получаем требуемый результат.]

10.23. Задача. Пусть $X_n(\omega)\geqslant X(\omega)$, $n=1, 2, \ldots$ ($X_n(\omega)\leqslant X(\omega)$, $n=1, 2, \ldots$) — измеримые функции, $X(\omega)$ — интегрируема. Вывести из 10.22, что

$$\int \underline{\lim}_{n\to\infty} X_n(\omega)\,\mu(d\omega)\leqslant \underline{\lim}_{n\to\infty}\int X_n(\omega)\,\mu(d\omega),$$

$$\left(\int \overline{\lim_{n\to\infty}}\, X_n(\omega)\,\mu(d\omega)\geqslant \overline{\lim_{n\to\infty}}\int X_n(\omega)\,\mu(d\omega)\right).$$

[Перейти к последовательности $Y_n(\omega)=X_n(\omega)-X(\omega)$ ($Y_n(\omega)=X(\omega)-X_n(\omega)$).]

10.24. Задача (теорема Лебега о мажорированной сходимости). Пусть $X_n(\omega)$, $n=1, 2, \ldots$, — сходящаяся последовательность измеримых функций, $|X_n(\omega)| \leqslant X(\omega)$, где $X(\omega)$ — интегрируема. Показать, что существует и конечен

$$\lim_{n\to\infty} \int X_n(\omega)\mu(d\omega) = \int \lim_{n\to\infty} X_n(\omega)\,\mu(d\omega).$$

[Ввиду разложения $X_n(\omega) = X_n^+(\omega) - X_n^-(\omega)$ достаточно разобрать случай неотрицательных $X_n(\omega)$, $n=1, 2, \ldots$. В этом случае, применяя 10.22, 10.23, получаем

$$\int \lim_{n\to\infty} X_n(\omega)\mu(d\omega) \leqslant \varliminf_{n\to\infty} \int X_n(\omega)\mu(d\omega) \leqslant$$

$$\leqslant \varlimsup_{n\to\infty} \int X_n(\omega)\mu(d\omega) \leqslant \int \lim_{n\to\infty} X_n(\omega)\mu(d\omega),$$

что и дает нужный результат.]

10.25. З а д а ч а. Математическое ожидание сл. в. $X(\omega)$ — измеримой функции на пространстве $(\Omega, \mathscr{A})$ с вероятностной мерой μ, в § 7 было определено как предел

$$\lim_{n\to\infty} \sum_{k=-\infty}^{\infty} kn^{-1}\mu(\omega : kn^{-1} < X(\omega) \leqslant (k+1)n^{-1})$$

в предположении, что ряды абсолютно сходятся (предел при этом существует, см. 7.17). Показать, что математическое ожидание существует и конечно тогда и только тогда, когда $X(\omega)$ интегрируема по мере μ, при этом

$$\mathbf{M}X = \int X(\omega)\mu(d\omega).$$

(Эта формула может служить определением математического ожидания и когда интеграл бесконечен.)

10.26. З а д а ч а. Пусть $X(\omega)$ измерима и $\mu(\omega : X(\omega) \neq 0) = 0$. Показать, что $\int X(\omega)\mu(d\omega) = 0$.

[Достаточно рассмотреть неотрицательную $X(\omega)$. Выбрав неубывающую последовательность простых функций $h_n(\omega)$, сходящуюся к $X(\omega)$, таким образом, что $h_n(\omega) = 0$ при $X(\omega) = 0$, получаем требуемое.]

10.27. З а д а ч а. Пусть $X(\omega)$, $Y(\omega)$ измеримы и $\mu(\omega : X(\omega) \neq$ $\neq Y(\omega)) = 0$. Показать, что

$$\int X(\omega)\mu(d\omega) = \int Y(\omega)\mu(d\omega),$$

если хотя бы один из интегралов существует.

[Достаточно рассмотреть неотрицательные X, Y. Если в этом случае оба интеграла бесконечны, то утверждение тривиально. В противном случае применяем 10.18 и 10.26.]

В теории вероятностей любое подмножество элементарных событий вероятности нуль пренебрежимо с точки зрения интерпретации модели. Точно так же случайная величина представляет интерес лишь с точностью до ее изменения на любом фиксированном множестве вероятности нуль. В теории интеграла Лебега мы видели, что измеримые функции, отличающиеся на множестве μ-нулевой меры, имеют одинаковые интегралы, а все свойства интеграла сохраняются, если соответствующие требования на функции накладывать не на всем множестве Ω, а на любом его подмножестве, дополнение к которому имеет μ-меру нуль. Например, результат 10.24 о переходе к пределу под знаком интеграла для мажорированной последовательности можно переформулировать в виде: если $\lim_{n\to\infty} X_n(\omega)=X(\omega)$, $|X(\omega)|\leqslant Y(\omega)$ всюду, за исключением некоторого множества меры нуль, а $Y(\omega)$ интегрируема, то

$$\lim_{n\to\infty}\int X_n(\omega)\,\mu(d\omega)=\int X(\omega)\,\mu(d\omega).$$

О свойствах, которые выполняются всюду, кроме, быть может, некоторого множества меры нуль, говорят как о свойствах, выполняющихся *почти всюду* (п. в.), а в теории вероятностей употребляют еще термин *«почти наверное»* (п. н.). Таким образом, можно говорить о п. в.-сходимости, если она имеет место всюду, кроме некоторого множества меры нуль, о п. в.-конечности и т. д. Более того, удобно рассматривать функции $X(\omega)$ на $(\Omega, \mathscr{A}, \mu)$, определенные всюду, кроме некоторого множества μ-меры нуль, и такие, что существует измеримая функция $\tilde{X}(\omega)$, что $\mu(\omega: X(\omega)\neq\tilde{X}(\omega))=0$. Под интегралом от такой μ-п. в. измеримой функции $X(\omega)$ подразумеваем интеграл от $\tilde{X}(\omega)$.

10.28. Задача. Пусть $X(\omega)\geqslant 0$, $\int X(\omega)\mu(d\omega)=0$. Показать, что $\mu(\omega: X(\omega)>0)=0$.

[Предположим, что $\mu(\omega: X(\omega)>0)>0$. По свойству непрерывности меры при $n\to\infty$

$$\mu(\omega: X(\omega)>n^{-1})\to\mu(\omega: X(\omega)>0),$$

так что, начиная с некоторого n, имеем $\mu(\omega: X(\omega)>n^{-1})>0$,

$$\int X(\omega)\mu(d\omega)\geqslant\int\limits_{\{\omega: X(\omega)>n^{-1}\}} X(\omega)\mu(d\omega)>0.$$

Полученное противоречие доказывает утверждение.]

10.29. Пример (интегралы Римана, Лебега, Римана — Стильтьеса, Лебега — Стильтьеса). Измеримая функция $g(x)$ на пространстве $(R, \mathscr{B})$, где $\mathscr{B}$ — класс одномерных борелевских множеств, называется также *борелевской*. Класс борелевских функций включает все обычно рассматриваемые в классическом анализе функции. Например, борелевскими являются непрерывные функции (как пределы простых), суперпозиция борелевских функций $(\{x: f(g(x))\in B\}=\{x: g(x)\in f^{-1}(B)\}\in\mathscr{B}$ для $B\in\mathscr{B})$.

Простая функция вида

$$h(x)=\sum_{i=1}^{n} I_{(t_{i-1},t_i]}(x),$$

где $a\equiv t_0<t_1<\ldots<t_n\equiv b$, $a_i\in R$, $i=1,\ \ldots,\ n$, называется *ступенчатой*. Интегралы Римана и Лебега от ступенчатых функций, очевидно, совпадают. Допустим, что функция $f(x)$ интегрируема по Риману на отрезке $[a,\ b]$. По критерию интегрируемости найдутся две последовательности $h_n'(x)$, $h_n''(x)$, $n=1,\ 2,\ \ldots$, ступенчатых функций, таких, что

$$h_n'(x)\leqslant f(x)\leqslant h_n''(x),\ \ x\in[a,\ b],\ \ h_n'(x)=h_n''(x)=0,\ \ x\overline{\in}[a,\ b],$$

$$\lim_{n\to\infty}\int h_n'(x)\,dx=\lim_{n\to\infty}\int h_n''(x)\,dx=\int_a^b f(x)\,dx,$$

где справа стоит интеграл Римана. Последовательности

$$\tilde{h}_n'(x)=\max_{k\leqslant n} h_k'(x),\ \ \tilde{h}_n''(x)=\min_{k\leqslant n} h_k''(x)$$

монотонны и обладают теми же свойствами, что и исходные. В таком случае функции

$$f'(x)=\lim_{n\to\infty}\tilde{h}_n'(x),\ \ f''(x)=\lim_{n\to\infty}\tilde{h}_n''(x)$$

измеримы, а их интегралы Лебега равны:

$$\int_{[a,b]} f'(x)\,dx=\lim_{n\to\infty}\int \tilde{h}_n'(x)\,dx=\lim_{n\to\infty}\int \tilde{h}_n''(x)\,dx=\int_{[a,b]} f''(x)\,dx.$$

Из соотношений

$$f'(x)\leqslant f(x)\leqslant f''(x),\ \ x\in[a,\ b],\ \ \int_{[a,b]}(f''(x)-f'(x))\,dx=0$$

вытекает (см. 10.28), что $\mu\{x:f''(x)-f'(x)\neq 0\}=0$. Таким образом, для интегрируемой по Риману на отрезке $[a,\ b]$ функции $f(x)$ построены измеримые функции $f'(x)$, $f''(x)$, которые отличаются от $f(x)$ на множестве лебеговой меры нуль, а интегралы Римана и Лебега для них равны:

$$\int_{[a,b]} f(x)\,dx=\int_{[a,b]} f'(x)\,dx=\int_{[a,b]} f''(x)\,dx.$$

Если функция $f(x)$ сама измерима, то интегралы Римана и Лебега от нее равны.

Измеримая функция $X(\omega)$ из $(\Omega, \mathscr{A}, \mu)$ в $(R, \mathscr{B})$ порождает на $\mathscr{B}$ меру $\mu_X(B)=\mu(\omega : X(\omega)\in B)$. Пусть $g(x)$ — борелевская функция. Покажем, что

$$\int g(X(\omega))\,\mu(d\omega)=\int g(x)\,\mu_X(dx),$$

где справа стоит интеграл Лебега на пространстве $(R, \mathscr{B}, \mu_X)$. Для $g(x)=I_B(x)$, $B\in\mathscr{B}$ имеем

$$\int I_B(X(\omega))\,\mu(d\omega)=\int\limits_{\{\omega: X(\omega)\in B\}} 1\cdot\mu(d\omega)=$$

$$=\mu(\omega : X(\omega)\in B)=\mu_X(B)=\int I_B(x)\,\mu_X(dx).$$

Так как простые функции — линейные комбинации индикаторов, утверждение верно для простых функций. Если $g(x)$ — любая неотрицательная борелевская функция и $h_n(x)$, $n=1, 2, \dots,$ — неубывающая последовательность неотрицательных простых функций, сходящаяся к $g(x)$, то в равенстве

$$\int h_n(X(\omega))\,\mu(d\omega)=\int h_n(x)\,\mu_X(dx)$$

можно перейти к пределу под знаком интеграла, воспользовавшись утверждением 10.15, и получить нужное равенство. Наконец, в общем случае положим $g(x)=g^+(x)-g^-(x)$ и воспользуемся уже установленным результатом для неотрицательных функций.

Применяя к вероятностным мерам, получаем важную формулу

$$\mathbf{M}g(X(\omega))=\int g(x)\,\mathbf{P}_X(dx),\quad \mathbf{P}_X(B)=\mathbf{P}(\omega : X(\omega)\in B),\quad B\in\mathscr{B}.$$

Любая σ-конечная мера μ на $(R, \mathscr{B})$ задается своей функцией распределения (ф. р.) $F(x)$ (см. 6.20). В частности, мере Лебега соответствует ф. р. $F(x)=x$. Теория интегрирования по Риману легко распространяется на произвольную ф. р. $F(x)$, если для ступенчатой функции

$$h(x)=\sum_{i=1}^{n} a_i I_{(t_i, t_{i+1}]}(x),\quad a\equiv t_0<t_1<\dots<t_n\equiv b,$$

положить

$$\int\limits_{a+}^{b} h(x)\,dF(x)=\sum_{i=1}^{n} a_i\,(F(t_i)-F(t_{i-1})).$$

Соответствующий интеграл называется *интегралом Римана — Стильтьеса*. Интегрируемые по Риману — Стильтьесу измеримые функции также интегрируемы по Лебегу по мере μ с ф. р. $F(x)$, а интегралы совпадают:

$$\int_{a+}^{b} g(x)\,dF(x)=\int_{(a,b]} g(x)\,\mu(dx).$$

Интеграл справа называют иногда *интегралом Лебега — Стильтьеса.*

10.30. Задача. Показать, что измеримая абсолютно интегрируемая по Риману на прямой функции также интегрируема по Лебегу и интегралы равны.

[Перейти к пределу в равенстве при n, $m\to\infty$:

$$\int_{-n}^{m} g(x)\,dx=\int g_{n,m}(x)\,dx,\quad g_{n,m}(x)=g(x)\,I_{[-n,m]}(x),$$

где слева стоит интеграл Римана, а справа — Лебега.]

Пусть $(\Omega_i, \mathscr{A}_i, \mu_i)$, $i=1, 2$, — два измеримых пространства с σ-конечными мерами. Пространство $(\Omega, \mathscr{A}, \mu)$, где $\Omega=\Omega_1\times\Omega_2$, $\mathscr{A}=\sigma(\mathscr{A}_1\times\mathscr{A}_2)$, μ — мера на $\mathscr{A}$, такая, что $\mu(A_1\times A_2)=$ $=\mu_1(A_1)\cdot\mu_2(A_2)$ для любых $A_i\in\mathscr{A}_i$, $i=1, 2$, называется произведением этих пространств с мерой. Существование меры μ вытекает из 6.26, где рассматриваются вероятностные меры μ_i, $i=1, 2$, если разбить Ω в счетное объединение непересекающихся прямоугольников $A_i^{(1)}\times A_j^{(2)}$, таких, что $\mu_1(A_i^{(1)})\cdot\mu_2(A_j^{(2)})<\infty$ (см. теорему о продолжении σ-конечной меры в § 6). Теория интегрирования по Лебегу дает возможность выразить в виде явной формулы меру μ на произведении пространств через меры μ_1, μ_2.

10.31.* Задача. Пусть $(\Omega, \mathscr{A})$ — произведение измеримых пространств $(\Omega_i, \mathscr{A}_i)$, $i=1, 2$. Для любого множества $A\subseteq\Omega$ и любой точки $\omega_1\in\Omega_1$ $(\omega_1\in\Omega_2)$ определим *сечение* $A_{\omega_1}=$ $=\{\omega_2:(\omega_1, \omega_2)\in A\}$ $(A_{\omega_2}=\{\omega_1:(\omega_1, \omega_2)\in A\})$ множества A координатой $\omega_1(\omega_2)$. Показать, что если $A\in\mathscr{A}$, то $A_{\omega_1}\in\mathscr{A}_1$ $(A_{\omega_2}\in\mathscr{A}_2)$; коротко: сечение измеримого множества измеримо.

[Легко видеть, что для любых A, $A^{(i)}\subseteq\Omega$, $i=1, 2, \ldots$,

$$(\bar{A})_{\omega_1}=\overline{A_{\omega_1}},\quad \Big(\bigcup_{(i)} A^{(i)}\Big)_{\omega_1}=\bigcup_{(i)} A^{(i)}_{\omega_1}.$$

Введем класс $\mathscr{A}'$ таких множеств $A\subseteq\Omega$, что $A_{\omega_1}\in\mathscr{A}_1$ при всех $\omega_1\in\Omega_1$. Если $A\in\mathscr{A}'$, то $A_{\omega_1}\in\mathscr{A}_1$, и поскольку $\mathscr{A}_1$ является σ-алгеброй, $\bar{A}_{\omega_1}\in\mathscr{A}_1$, а из предыдущего вытекает, что $\bar{A})_{\omega_1}\in\mathscr{A}_1$, т. е. $\bar{A}'\in\mathscr{A}'$. Точно так же убеждаемся, что $\mathscr{A}'$ замкнут относительно счетных объединений. Очевидно, $\varnothing$, $\Omega\in\mathscr{A}'$, так что $\mathscr{A}'$ является σ-алгеброй. Поскольку для любого прямоугольника $A=A_1\times A_2$ имеем $A_{\omega_1}=A_2$ при $\omega_1\in A_1$ и $A_{\omega_1}=\varnothing$ при $\omega_1\bar{\in} A_1$, то все прямоугольники со сторонами $A_i\in\mathscr{A}_i$, $i=1, 2$, входят в $\mathscr{A}'$. Учитывая, что $\mathscr{A}$ есть минимальная σ-алгебра, содержащая измеримые прямоугольники, заключаем, что $\mathscr{A}'\supseteq\mathscr{A}$.]

Ниже во избежание недоразумений измеримые на $(\Omega, \mathscr{A})$ функции называются также *измеримыми относительно* $\mathscr{A}$, или $\mathscr{A}$*-измеримыми.*

10.32. Задача. В условии 10.31 для фиксированного $\omega_1 \in \Omega_1$ и любой функции $f(\omega)$, $\omega \in \Omega$, определим сечение $f_{\omega_1}(\omega_2)$ как функцию на Ω_2, задаваемую равенством $f_{\omega_1}(\omega_2) = f(\omega_1, \omega_2)$. Показать, что если $f(\omega)$ измерима относительно $\mathscr{A}$, то $f_{\omega_1}(\omega_2)$ измерима относительно $\mathscr{A}_2$.

[Вытекает из равенства

$$f_{\omega_1}^{-1}(B) = \{\omega_2 : f_{\omega_1}(\omega_2) \in B\} = \{\omega_2 : f(\omega_1, \omega_2) \in B\} =$$
$$= \{\omega_2 : (\omega_1, \omega_2) \in f^{-1}(B)\} = (f^{-1}(B))_{\omega_1}$$

и утверждения 10.31.]

10.33.* Задача. Пусть $(\Omega, \mathscr{A}, \mu)$ — произведение измеримых пространств $(\Omega_i, \mathscr{A}_i, \mu_i)$ с вероятностными мерами μ_i, $i=1, 2$. Введем класс $\mathscr{M}$ множеств $A \subseteq \Omega$, для которых функция $\mu_1(A_{\omega_2})$ $\mathscr{A}_2$-измерима и справедливо равенство

$$\mu(A) = \int \mu_1(A_{\omega_2})\, \mu_2(d\omega_2).$$

Показать, что $\mathscr{M}$ является монотонным классом и содержит алгебру конечных объединений прямоугольников $A_1 \times A_2$, $A_i \in \mathscr{A}_i$, $i=1, 2$, а следовательно, содержит $\mathscr{A}$ (см. 6.36). Распространить утверждение на σ-конечные меры μ_i, $i=1, 2$.

[Если $A = A_1 \times A_2$, $A_i \in \mathscr{A}_i$, $i=1, 2$, то функция

$$\mu_1(A_{\omega_2}) = \mu_1(A_1)\, I_{A_2}(\omega_2)$$

— простая на $(\Omega_2, \mathscr{A}_2)$. Интеграл от нее по мере μ_2 равен $\mu_1(A_1) \cdot \mu_2(A_2) \equiv \mu(A)$, так что $A \in \mathscr{M}$, т. е. класс $\mathscr{M}$ содержит полуалгебру прямоугольников $A_1 \times A_2$, $A_i \in \mathscr{A}_i$, $i=1, 2$ (см. 6.24).

Если прямоугольники $A^{(j)} = A_1^{(j)} \times A_2^{(j)}$, $j = 1, \ldots, n$, попарно не пересекаются, $A = \bigcup_{(j)} A^{(j)}$, то, пользуясь доказанным, получаем

$$\mu(A) = \sum_{(j)} \mu_1(A_1^{(j)})\, \mu_2(A_2^{(j)}) = \int \sum_{(j)} \mu_1(A_{\omega_2}^{(j)})\, I_{A_2^{(j)}}(\omega_2)\, \mu_2(d\omega_2).$$

Так как $A_{\omega_2} = \bigcup_{(j)} A_{\omega_2}^{(j)}$, причем множества $A_{\omega_2}^{(j)}$, $j = 1, \ldots, n$, попарно не пересекаются, то

$$\sum_{(j)} \mu_1(A_{\omega_2}^{(j)})\, I_{A_2^{(j)}}(\omega_2) = \mu_1(A_{\omega_2})\, I_{A_2}(\omega_2)$$

и, следовательно, $A \in \mathscr{M}$, т. е. $\mathscr{M}$ содержит алгебру конечных объединений прямоугольников (см. 6.1).

Пусть теперь $A^{(1)}\subseteq A^{(2)}\subseteq,\ldots, A^{(n)}\in\mathcal{M}$, $n=1, 2, \ldots$. Переходя к пределу при $n\to\infty$ в равенствах

$$\mu(A^{(n)})=\int\mu_1(A^{(n)}_{\omega_2})\mu_2(d\omega_2),\quad n=1, 2, \ldots,$$

и пользуясь 10.15 и свойством непрерывности меры μ_1, получаем, что $\bigcup\limits_{(n)}A^{(n)}\in\mathcal{M}$. Используя 10.20 вместо 10.15, заключаем точно так же, что класс $\mathcal{M}$ замкнут относительно пересечений убывающих последовательностей (можно также воспользоваться равенством $\mu(\bar{A}^{(n)})=1-\mu(A^{(n)})$). В случае σ-конечных мер достаточно разбить пространство в счетное объединение прямоугольников с конечными мерами сторон.]

10.34.* Задача. В условии 10.33: (I) пусть интеграл от функции $f(\omega)=f(\omega_1, \omega_2)$ существует, показать, что

$$\int f(\omega)\mu(d\omega)=\int\left(\int f_{\omega_2}(\omega_1)\mu_1(d\omega_1)\right)\mu_2(d\omega_2);$$

(II) интеграл слева существует, если

$$\int\left(\int|f_{\omega_2}(\omega_1)|\mu_1(d\omega_1)\right)\mu_2(d\omega_2)<\infty.$$

[(I) Для $f(\omega)=I_A(\omega)$, $A\in\mathcal{A}$, утверждение сводится к 10.33, при этом

$$\int f_{\omega_2}(\omega_1)\mu_1(d\omega_1)=\mu_1(A_{\omega_2})$$

является $\mathcal{A}_2$-измеримой функцией. Поскольку для любых функций f, g и постоянных a, b имеем, очевидно,

$$(af+bg)_{\omega_2}=af_{\omega_2}+bg_{\omega_2},$$

то утверждение сохраняется и для простых функций — линейных комбинаций индикаторов. Для любой неотрицательной измеримой функции $f(\omega)$ рассмотрим неубывающую последовательность сходящихся к ней неотрицательных простых функций $f^{(n)}(\omega)$, $n=1, 2, \ldots$, и запишем для них уже установленные равенства

$$\int f^{(n)}(\omega)\mu(d\omega)=\int\left(\int f^{(n)}_{\omega_2}(\omega_1)\mu_1(d\omega_1)\right)\mu_2(d\omega_2),$$

где внутренний интеграл справа — измеримая функция на $(\Omega_2, \mathcal{A}_2)$. Так как последовательность $f^{(n)}_{\omega_2}(\omega_1)$, $n=1, 2, \ldots$, не убывает, то не убывает последовательность

$$\int f^{(n)}_{\omega_2}(\omega_1)\mu_1(d\omega_1),\quad n=1, 2, \ldots .$$

Переходя в написанном выше равенстве к пределу при $n\to\infty$ и пользуясь определением интеграла от $f(\omega)$ и теоремой о монотонной сходимости 10.15, получаем требуемое утверждение для случая неотрицательной $f(\omega)$. При этом, если интеграл от $f(\omega)$ конечен, то функция

$$\int f_{\omega_2}(\omega_1)\,\mu_1(d\omega_1)$$

конечна почти всюду относительно μ_2.

В общем случае из разложения $f(\omega)=f^+(\omega)-f^-(\omega)$ имеем

$$\int f(\omega)\,\mu(d\omega)=\int f^+(\omega)\,\mu(d\omega)-\int f^-(\omega)\,\mu(d\omega)=$$

$$=\int\left(\int f^+_{\omega_2}(\omega_1)\,\mu_1(d\omega_1)\right)\mu_2(d\omega_2)-\int\left(\int f^-_{\omega_2}(\omega_1)\,\mu_1(d\omega_1)\right)\mu_2(d\omega_2),$$

и остается убедиться, что выражение

$$\int f^+_{\omega_2}(\omega_1)\,\mu_1(d\omega_1)-\int f^-_{\omega_2}(\omega_1)\,\mu_1(d\omega_1)$$

обращается в неопределенность $\infty-\infty$ не более чем на некотором множестве μ_2-меры нуль. Это, однако, вытекает из того, что хотя бы один из интегралов от $f^+(\omega)$ или $f^-(\omega)$ конечен.

(II) Ввиду представлений $f=f^+-f^-$, $|f|=f^++f^-$ достаточно рассмотреть неотрицательную $f(\omega)$. Но в этом случае утверждение непосредственно следует из (I), так как интеграл от неотрицательной измеримой функции всегда существует.]

Интеграл Лебега от функции $f(\omega)=f(\omega_1, \omega_2)$ на произведении пространств $(\Omega_i, \mathscr{A}_i, \mu_i)$, $i=1, 2$, записывают в форме двойного интеграла, а утверждение 10.34, называемое *теоремой Фубини*, выражается равенством двойного и повторного интегралов:

$$\iint f(\omega_1, \omega_2)\,\mu_1(d\omega_2)\,\mu_2(d\omega_2)=\int\mu_2(d\omega_2)\int f(\omega_1, \omega_2)\,\mu_1(d\omega_1).$$

10.35.* Задача. Сл. в. $X(\omega)$ и $Y(\omega)$ независимы, т. е. для любых промежутков $I_{a,b}$, $I_{c,d}$ выполняется равенство

$$\mathbf{P}_{X,Y}(I_{a,b}\times I_{c,d})=\mathbf{P}_X(I_{a,b})\,\mathbf{P}_Y(I_{c,d}).$$

Показать: (I) что мера $\mathbf{P}_{X,Y}(B)=\mathbf{P}(\omega:(X(\omega), Y(\omega))\in B)$ определена этим равенством на классе $\mathscr{B}^2$ борелевских множеств на плоскости; (II) что пространство $(R^2, \mathscr{B}^2, \mathbf{P}_{X,Y})$ — произведение пространств $(R, \mathscr{B}, \mathbf{P}_X)$ и $(R, \mathscr{B}, \mathbf{P}_Y)$.

[(I) Обозначим $\mathscr{B}'$ класс всех таких множеств $B'\in R^2$, что множество $\{\omega:(X(\omega), Y(\omega))\in B'\}$ — измеримое. Очевидно, что для любых B_1, $B_2\in\mathscr{B}$ имеем $B_1\times B_2\in\mathscr{B}'$. Если $B'\in\mathscr{B}'$, то

$$\{\omega:(X(\omega), Y(\omega))\in\bar{B}'\}=\overline{\{\omega:(X(\omega), Y(\omega))\in B'\}}$$

и $\bar{B}'\in\mathscr{B}$. Аналогично, если $B_i'\in\mathscr{B}'$, $i=1, 2, \ldots$, то $\bigcup\limits_{\{i\}}B_i'\in\mathscr{B}'$. Следовательно, $\mathscr{B}'\supseteq\mathscr{B}^2$ и мера $\mathbf{P}_{X,Y}$ определена на $\mathscr{B}^2$. (II) Класс множеств B_1, таких, что при заданном промежутке I верно равенство

$$\mathbf{P}_{X,Y}(B_1\times I)=\mathbf{P}_X(B_1)\,\mathbf{P}_Y(I),$$

является, очевидно, σ-алгеброй и содержит все промежутки, а следовательно, и класс борелевских множеств. Фиксируя произвольное $B_1 \in \mathscr{B}$, проводим такое же рассуждение для класса множеств B_2, для которых

$$\mathbf{P}_{X,Y}(B_1 \times B_2) = \mathbf{P}_X(B_1)\,\mathbf{P}_Y(B_2),$$

и получаем, что равенство верно при всех B_1, $B_2 \in \mathscr{B}$.]

10.36. Задача. Сл. в. X_1, X_2 независимы, а их математические ожидания конечны. Показать, что $\mathbf{M}X_1X_2 = \mathbf{M}X_1 \cdot \mathbf{M}X_2$.
[Вытекает из 10.34, 10.35.]

10.37. Задача. Сл. в. X, Y независимы. Показать, что для любого борелевского множества B

$$\mathbf{P}_{X+Y}(B) = \int \mathbf{P}_X(B-y)\,\mathbf{P}_Y(dy) = \int \mathbf{P}_Y(B-x)\,\mathbf{P}_X(dx),$$

где $B-a = \{x-a : x \in B\}$.
[Применяя 10.34, имеем

$$\mathbf{P}_{X+Y}(B) = \int\int I_B(x+y)\,\mathbf{P}_X(dx)\,\mathbf{P}_Y(dy) =$$
$$= \int \mathbf{P}_Y(dy) \int I_B(x+y)\,\mathbf{P}_X(dx) = \int \mathbf{P}_X(B-y)\,\mathbf{P}_Y(dy).]$$

10.38. Задача. Пусть сл. в. X_1, X_2 имеют совместную плотность $f(x_1, x_2)$, т. е.

$$\mathbf{P}(\omega: (X_1(\omega), X_2(\omega)) \in B) = \int\int_B f(x_1, x_2)\,dx_1 dx_2$$

для любого двумерного борелевского множества B. Показать, что X_1 (X_2) имеет плотность

$$f_{X_1}(x_1) = \int f(x_1, x_2)\,dx_2 \quad \left(f_{X_2}(x_2) = \int f(x_1, x_2)\,dx_1\right).$$

Обобщить утверждение на любой набор $X_1(\omega), \ldots, X_n(\omega)$ случайных величин.
[Для $B = B_1 \times R$, где B_1 — линейное борелевское множество, получаем, применяя 10.34,

$$\mathbf{P}(\omega: X_1(\omega) \in B_1) = \mathbf{P}(\omega: (X_1(\omega), X_2(\omega)) \in B) =$$
$$= \int\int I_B(x_1, x_2) f(x_1, x_2)\,dx_1\,dx_2 = \int I_{B_1}(x_1)\,dx_1 \int f(x_1, x_2)\,dx_2 =$$
$$= \int_{B_1} f_{X_1}(x_1)\,dx_1, \quad f_{X_1}(x_1) = \int f(x_1, x_2)\,dx_2.$$

Отметим, что плотность распределения вероятностей определена с точностью до ее изменения на произвольном множестве меры Лебега нуль, и, в частности, можно считать, что плотность задана μ-почти всюду.]

10.39. Задача. Пусть мера μ в $(R^n, \mathscr{B}^n)$ задается плотностью $f(x) \geqslant 0$:

$$\mu(B) = \int_B f(x)\,dx, \quad B \in \mathscr{B}^n.$$

Тогда для любой измеримой функции $h(x)$

$$\int h(x)\mu(dx)=\int h(x)f(x)\,dx$$

в предположении, что один из интегралов существует.

[Для $h(x)=I_B(x)$ равенство сводится к определению меры μ. Отсюда следует, что оно справедливо для простых функций. Для неотрицательных функций оно получается переходом к пределу по неубывающей последовательности простых неотрицательных функций, сходящихся к $h(x)$. Общий случай получается из разложения $h=h^+-h^-$.]

10.40. Задача. Сл. в. $X(\omega)$, $Y(\omega)$ независимы и имеют плотности. Тогда сл. в. $X(\omega)$, $Y(\omega)$ имеют совместную плотность $f_{X,Y}(x, y)=f_X(x)f_Y(y)$, а сл. в. $X+Y$ имеет плотность

$$f_{X+Y}(u)=\int f_X(u-y)f_Y(y)\,dy.$$

[Применяя 10.34, 10.39, имеем для любого $B\in\mathscr{B}^2$

$$\mathbf{P}_{X,Y}(B)=\int \mathbf{P}_Y(dy)\int I_B(x, y)\mathbf{P}_X(dx)=$$

$$=\int f_Y(y)\,dy\int I_B(x, y)f_X(x)\,dx=\iint\limits_B f_X(x)f_Y(y)\,dx\,dy,$$

что устанавливает первое из утверждений 10.40. Второе устанавливается аналогично.]

10.41. Пример (марковские процессы). В 4.27 на эвристическом уровне вводился процесс броуновского движения $W_t\equiv W(t)$, $t\geqslant 0$, как семейство сл. в., для которых при любых $0<t_1<\ldots<t_n$, $n=1, 2, \ldots$, совместная плотность сл. в. $W(t_1), \ldots, W(t_n)$ имеет вид

$$f_{t_1,\ldots,t_n}(x_1, \ldots, x_n)=\prod_{i=1}^{n} p(t_{i-1}, x_{i-1}, t_i, x_i), \quad t_0=x_0=0,$$

$$p(s, x, t, y)=\left(1/\sqrt{t-s}\right)\varphi\left((y-x)/\sqrt{t-s}\right), \quad 0\leqslant s<t,$$

где $\varphi(x)$ — стандартная нормальная плотность. Строгая теория требует построения вероятностного пространства, на котором такое семейство могло бы быть задано. Это построение не зависит от частного вида функции $p(s, x, t, y)$ при условии, что она является плотностью распределения вероятностей по y, $\mathscr{B}^2$-измерима по паре переменных (x, y) и удовлетворяет *уравнениям Колмогорова—Чепмена:*

$$\int\limits_{-\infty}^{\infty} p(s, x, t, y)p(t, y, u, z)\,dy=p(s, x, u, z), \quad s<t<u.$$

Заметим, что в частном случае броуновского движения левая часть

$$\int_{-\infty}^{\infty} (1/\sqrt{t-s})\,\varphi((y-x)/\sqrt{t-s})\,(1/\sqrt{u-t})\,\varphi((z-y)/\sqrt{u-t})\,dy$$

представляет собой свертку двух нормальных плотностей с параметрами x, $\sqrt{t-s}$ и 0, $\sqrt{u-t}$ а потому является нормальной плотностью с параметрами x, $\sqrt{u-s}$, т. е. равна $(1/\sqrt{u-s})\,\varphi((z-$ $-x)/\sqrt{u-s})$, и, следовательно, уравнения Колмогорова—Чепмена удовлетворяются.

Введем множество $\Omega=\{\omega=x(t),\ t\geqslant 0\}$ действительных функций на $[0, \infty)$ и класс $\mathcal{U}$ *цилиндрических* подмножеств

$$B_{t_1,\ldots,t_n}=\{\omega:(x(t_1),\ \ldots,\ x(t_n))\in B^{(n)}\},$$

где $0<t_1<\ldots<t_n$, $n=1,\ 2,\ \ldots,$ $B^{(n)}\in\mathcal{B}^n$, $\mathcal{B}^n$ — σ-алгебра n-мерных борелевских множеств. Обозначив $\mathbf{P}_{t_1,\ldots,t_n}(B^{(n)})$, $B^{(n)}\in\mathcal{B}^n$, меру с плотностью $f_{t_1,\ldots,t_n}(x_1,\ \ldots,\ x_n)$, определим на $\mathcal{U}$ функцию $\mathbf{P}$, полагая

$$\mathbf{P}(B_{t_1,\ldots,t_n})=\mathbf{P}_{t_1,\ldots,t_n}(B^{(n)}).$$

Класс $\mathcal{U}$ — алгебра множеств. Из уравнений Колмогорова—Чепмана легко следует, что выполняются так называемые *условия согласованности* (ср. 6.29):

$$\mathbf{P}_{t_1,\ldots,t_n}(B_1,\ \ldots,\ B_{i-1},\ R,\ B_{i+1},\ \ldots,\ B_n)=$$

$$=\mathbf{P}_{t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_n}(B_1,\ \ldots,\ B_{i-1},\ B_{i+1},\ \ldots,\ B_n)$$

при любых $0<t_1<\ldots<t_n$, $B_i\in\mathcal{B}$, $i=1,\ldots,n$, благодаря чему функция $\mathbf{P}$ на $\mathcal{U}$ определена корректно. Очевидно, $\mathbf{P}$ — конечно-аддитивная вероятностная мера на $\mathcal{U}$. Доказательство непрерывности меры $\mathbf{P}$ в $\varnothing$ проводится по уже использованной в 6.27, 6.29 схеме. Тем самым мера $\mathbf{P}$ продолжается на $\sigma(\mathcal{U})$.

На вероятностном пространстве $(\Omega,\ \sigma(\mathcal{U}),\ \mathbf{P})$ можно задать непосредственно семейство сл. в. $X_t(\omega)$, $t\geqslant 0$, полагая при каждом фиксированном $t'>0$ $X_{t'}(\omega)=x(t')$, где $\omega=x(t)$, $t\geqslant 0$ $(X_0(\omega)=$ $=0)$. Сл. в. $X_{t_1},\ \ldots,\ X_{t_n}$, $0<t_1<\ldots<t_n$, имеют совместную плотность распределения $f_{t_1,\ldots,t_n}$, а условная плотность

$$f_{X(t_n)|X(t_1),\ldots,X(t_{n-1})}(x_n|x_1,\ \ldots,\ x_{n-1})=p(t_{n-1},\ x_{n-1},\ t_n,\ x_n)$$

не зависит от t_i, x_i, $i<n-1$. Процесс с таким свойством называется *марковским,* функция $p(s,\ x,\ t,\ y)$ — его *переходной плотностью.* Если переходная плотность зависит от s, t через разность, то процесс называется *однородным*.

Глава III

НЕКОТОРЫЕ МОДЕЛИ И МЕТОДЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

§ 11.

ПРОСТОЕ СИММЕТРИЧНОЕ БЛУЖДАНИЕ

Простое симметричное блуждание конечной длительности было введено в § 1, где прямыми комбинаторными вычислениями были найдены вероятности ряда событий, описывающих поведение блуждающей частицы. Возможность получать точные формулы делает модель простого блуждания чрезвычайно полезной для исследования закономерностей, которые, как оказывается, справедливы и в значительно более общей ситуации. Эти закономерности имеют форму так называемых предельных теорем. С некоторыми частными случаями мы познакомились в § 3 (см. 3.16, 3.19). Модель неограниченного во времени блуждания, построенная в § 6, позволяет рассматривать вопросы, связанные с поведением всей бесконечной траектории движения, и, кроме того, делает теорию предельного поведения распределений, связанных с конечными отрезками блуждания, более прозрачной.

Отсылая за подробностями к § 6, считаем заданным вероятностное пространство $(\Omega, \mathcal{A}, \mathbf{P})$, на котором определена бесконечная последовательность независимых сл. в. $X_i = X_i(\omega)$, $i=1, 2, \ldots$, принимающих два значения 1 и -1 с равными вероятностями. Неограниченное случайное блуждание определим как последовательность сл. в. $S_n^{(a)} = a + X_1 + \ldots + X_n$, $n = 1, 2, \ldots, S_0^{(a)} = a$. Это блуждание начинается из (целой) точки a. Положим $S_n^{(0)} = S_n$. Для дальнейшего существенно, что алгебра событий $\mathcal{A}$ замкнута относительно образования счетных объединений и пересечений (является σ-алгеброй), а вероятностная мера $\mathbf{P}$ счетно-аддитивна:

$$\mathbf{P}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathbf{P}(A_i), \quad A_i \in \mathcal{A}, \quad A_i A_j = \varnothing, \quad i \neq j.$$

Непосредственным следствием счетной аддитивности меры $\mathbf{P}$ является свойство *непрерывности* меры (см. 6.7): для любой возрастающей (убывающей) последовательности событий $A_1 \subseteq A_2 \subseteq \ldots$

$\subseteq (A_1 \supseteq A_2 \supseteq \ldots)$ и $A = \bigcup_{i=1}^{\infty} A_i$ $\left(A = \bigcap_{i=1}^{\infty} A_i\right)$ имеет место соотношение

$$\lim_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}(A).$$

11.1. Задача. Пусть $T_1 = T_1(\omega)$ — момент первого достижения блуждающей частицей, выходящей из нуля, положения 1: $T_1(\omega) = k \Leftrightarrow S_1(\omega), \ldots, S_{k-1}(\omega) \leqslant 0$, $S_k(\omega) = 1$, $k = 1, 2, \ldots$, и $T_1(\omega) = \infty$, если $S_k(\omega) \leqslant 0$ при всех k. Используя формулу (см. 1.15)

$$\mathbf{P}(\omega : S_1(\omega) \leqslant 0, \ \ldots, \ S_{2n}(\omega) \leqslant 0) = C_{2n-1}^{n} 2^{-2n+1}$$

показать, что $\mathbf{P}(\omega : T_1(\omega) = \infty) = 0$.
[По свойству непрерывности вероятностной меры

$$\mathbf{P}(\omega : T_1(\omega) = \infty) = \lim_{n\to\infty} \mathbf{P}(\omega : S_1(\omega) \leqslant 0, \ \ldots, \ S_{2n}(\omega) \leqslant 0) = 0,$$

так как (см. 3.19)

$$C_{2n-1}^{n} 2^{-2n+1} \sim (\pi n)^{-1/2}, \ n \to \infty.]$$

11.2. Задача. Пусть $N_1 = N_1(\omega)$ — момент первого возвращения блуждающей частицы в начальное состояние, $N_1(\omega) = \infty$, если $S_k(\omega) \neq 0$ при всех k. Показать, что $\mathbf{P}(\omega : N_1(\omega) = \infty) = 0$.

[$\mathbf{P}(\omega : N_1(\omega) > 2n) = 2\mathbf{P}(\omega : S_k(\omega) > 0, \ k = 1, 2, \ldots, 2n) = 2C_{2n-1}^{n} 2^{-2n}$ (см. 1.14) и далее аналогично 11.1]

Сл. в. T_1, N_1 в 11.1, 11.2 — частный случай так называемых *моментов остановки*. Представим себе, что ведется наблюдение за некоторой последовательностью сл. в. $Z_1, Z_2, \ldots$. Предположим, что наблюдения ведутся до некоторого случайного момента времени $N \leqslant \infty$, причем решение об остановке делается на основании уже проведенных наблюдений. Формально такое правило можно описать последовательностью борелевских подмножеств $B_n \subseteq R^n$, таких, что

$$N(\omega) = n \Leftrightarrow (Z_1(\omega), \ldots, Z_n(\omega)) \in B_n, \ n = 1, 2, \ldots.$$

Так, определенная сл. в. $N(\omega)$ называется *моментом остановки для последовательности* Z_i, $i = 1, 2, \ldots$.

11.3. Задача. Пусть Z_i, $i = 1, 2, \ldots$, — последовательность независимых одинаково распределенных дискретных случайных величин, N — момент остановки, $\mathbf{P}(\omega : N(\omega) < \infty) = 1$. Показать, что последовательность $Y_k = Z_{N+k}$, $k = 1, 2, \ldots$, распределена так же, как исходная последовательность Z_i, $i = 1, 2, \ldots$, и не зависит от $(N, Z_1, \ldots, Z_N)$.

$$[\mathbf{P}(\omega : N(\omega) = n, \ Z_i(\omega) = z_i, \ i = 1, \ \ldots, \ n, \ Y_k(\omega) = y_k, \ k = 1, \ \ldots, \ m) =$$
$$= \mathbf{P}(\omega : N(\omega) = n, \ Z_i(\omega) = z_i, \ i = 1, \ \ldots, \ n, \ Z_{n+k}(\omega) = y_k, \ k = 1, \ \ldots, \ m) =$$
$$= \mathbf{P}(\omega : N(\omega) = n, \ Z_i(\omega) = z_i, \ i = 1, \ \ldots, \ n) \cdot \mathbf{P}(\omega : Z_{n+k}(\omega) = y_k,$$

$k=1, \ldots, m)=\mathbf{P}(\omega : N(\omega)=n,\ Z_i(\omega)=z_i,\ i=1, \ldots, n)\cdot\mathbf{P}(\omega : Z_k(\omega)=$
$=y_k,\ k=1, \ldots, m)$.

Суммируя по всем возможным наборам $(n, z_1, \ldots, z_n)$, получаем

$$\mathbf{P}(\omega : Y_k(\omega)=y_k,\ k=1, \ldots, m)=\mathbf{P}(\omega : Z_k(\omega)=y_k,\ k=1, \ldots, m),$$

что и заканчивает доказательство.]

11.4. Задача. В условии 11.3 положим $N_1(\omega)=N(\omega)$, определим $N_2(\omega)$ по последовательности Z_{N_1+k}, $k=1, 2, \ldots$, точно так же, как $N(\omega)$ определено по последовательности Z_i, $i=1, 2, \ldots$, аналогично определим $N_3(\omega)$ по последовательности $Z_{N_1+N_2+k}$, $k=1,2,\ldots$, и т. д. Показать, что отрезки $(N_{i+1}, Z_{N_i+1}, \ldots, Z_{N_i+N_{i+1}-1})$ $i=0, 1, 2, \ldots$, $N_0=0$, независимы и одинаково распределены.

Для первых двух отрезков выкладки аналогичны 11.3, далее можно использовать индукцию.]

11.5. Задача. Пусть $T_a=T_a(\omega)$ — момент первого достижения блуждающей частицей, выходящей из нуля, положения a. Показать, что: (I) последовательность S_{T_a+k}, $k=0, 1, 2, \ldots$, представляет собой простое симметричное блуждание, начинающееся из точки a и не зависящее от S_h, $k\leqslant T_a$; (II) отрезки $(S_k-a,\ T_a\leqslant k\leqslant \leqslant T_{a+1})$, $a=0, 1, 2, \ldots, T_0=0$, независимы и одинаково распределены.

[(I) Ввиду симметрии блуждания достаточно рассмотреть $a>0$. Поскольку сл. в. $X_1, \ldots, X_n$ и $S_1, \ldots, S_n$ выражаются однозначно друг через друга, то $T_1(\omega)$ — момент остановки как относительно последовательности S_i, $i=1, 2, \ldots$, так и относительно X_i, $i=1, 2, \ldots$. Тогда (см. 11.3) последовательность X_{T_1+k}, $k=1, 2, \ldots$, распределена так же, как X_i, $i=1, 2, \ldots$, и не зависит от X_i, $i\leqslant T_1$. Но $S_{T_1+k}=1+X_{T_1+1}+\ldots+X_{T_1+k}$, $k=1, 2, \ldots$, что устанавливает утверждение (I) при $a=1$. Общий случай выводится по индукции. Утверждение (II) легко следует из (I).]

11.6. Задача. Показать, что с вероятностью 1

$$\overline{\lim_{n\to\infty}}\, S_n(\omega)=\infty, \quad \underline{\lim_{n\to\infty}}\, S_n(\omega)=-\infty.$$

[Нетрудно видеть (см. 11.5), что при любом $a\neq 0$ сл. в. T_a представима как сумма $|a|$ независимых случайных величин, распределенных так же, как T_1. В частности, $\mathbf{P}(\omega : T_a(\omega)<\infty)=1$, а следовательно,

$$\mathbf{P}\Big(\bigcap_{a=-\infty}^{\infty}\{\omega : T_a(\omega)<\infty\}\Big)=1,$$

т. е. с вероятностью 1 траектория блуждания достигает любого состояния, что и дает нужный результат.]

11.7. Пример (предельная теорема для времен достижения). Ряд предельных соотношений для симметричного блуждания, полученных в § 3, имеет следующую форму. Дана сл. в. Z_n — функция от отрезка траектории $S_1, \ldots, S_n$. Ее распределение вероятностей $\mathbf{P}_{Z_n}(B)=\sum_{\{k\in B\}} p_{Z_n}(k)$ таково, что $p_{Z_n}(k)\to 0$ при $n\to\infty$ и любом фиксированном k. Подбирается масштабное преобразование прямой с коэффициентом $\lambda_n\to\infty$ при $n\to\infty$, так что для любого промежутка $I_{a,b}$ имеет место предельное соотношение

$$\mathbf{P}_{Z_n}(\lambda_n I_{a,b})=\mathbf{P}_{Z_n/\lambda_n}(I_{a,b})\to\mathbf{P}(I_{a,b}),\quad n\to\infty,$$

где $\mathbf{P}(B)$ — некоторая вероятностная мера на прямой.

Так, было установлено, что распределение сл. в. $S_n/\sqrt{n}$ сходится к стандартному нормальному:

$$\mathbf{P}\left(\omega: S_n(\omega)/\sqrt{n}\in I_{a,b}\right)\to\int_a^b\varphi(x)\,dx,\quad \varphi(x)=(2\pi)^{-1/2}e^{-x^2/2}.$$

Для сл. в. $M_n(\omega)=\max\{S_k(\omega),\ k\leqslant n\}$ из 1.17, 3.19 легко вывести следующую предельную теорему:

$$\mathbf{P}\left(\omega: M_n(\omega)/\sqrt{n}\in[0,\ x]\right)\to 2\int_0^x\varphi(t)\,dt,\quad n\to\infty.$$

Из этих соотношений можно заключить, что за время n блуждающая частица успевает побывать на расстоянии порядка $\sqrt{n}$ от начального положения.

Зададимся вопросом, сколько времени требуется частице для достижения уровня $a\gg 1$. Предыдущие рассуждения делают правдоподобным вывод, что при $a\to\infty$ это время имеет порядок a^2. Точная формула для распределения вероятностей сл. в. T_a вытекает из 1.13—1.16. При a четном имеем

$$\mathbf{P}(\omega: T_{2k}(\omega)=2m)=(C_{2m-1}^{m+k-1}-C_{2m-1}^{m+k})\,2^{-2m}=\frac{k}{m}C_{2m}^{m+k}2^{-2m},$$

при a нечетном вычисления аналогичны:

$$\mathbf{P}(\omega: T_{2k+1}(\omega)=2m+1)=\frac{2k+1}{2m+1}C_{2m+1}^{m+k+1}2^{-2m-1}.$$

Первое представление о распределении вероятностной массы можно получить, рассматривая отношение

$$\mathbf{P}(\omega: T_{2k}(\omega)=2m)/\mathbf{P}(\omega: T_{2k}(\omega)=2(m+1))=$$
$$=\frac{k(m+1)\,(2m)!\,(m+k+1)!\,(m-k+1)!}{mk\,(m+k)!\,(m-k)!\,(2m+2)!}\cdot 2^2=$$
$$=\frac{(m+1)(m+k+1)(m-k+1)}{m(2m+2)(2m+1)}\cdot 2^2=\frac{(1+1/m)(1-k^2/(m+1)^2)}{1-1/(2m+2)}.$$

При $m/k\to\infty$, пренебрегая остаточными членами, запишем его в виде $1+3/(2m)-k^2/m^2$. Это выражение обращается в 1 при $m=$

$=(2/3)k^2$, так что распределение вероятностей имеет максимум вблизи точки $m=[(2/3)k^2]$. В стороны от точки максимума распределение убывает.

Воспользовавшись 3.19, запишем следующую асимптотику для распределения вероятностей сл. в. T_{2k}:

$$(k/m)C_{2m}^{m+k}2^{-2m}\sim(k/m)(\pi m)^{-1/2}\exp(-k^2/m)$$

равномерно по таким m, что $mk^{-4/3}\to\infty$. В частности,

$$\mathbf{P}(\omega: T_{2k}(\omega)/(2k)^2=x)\sim(2/\pi)^{1/2}x^{-3/2}\exp(-1/(2x))\cdot(2k)^{-2}$$

равномерно по $x=x_{m,k}=2m/(2k)^2$, лежащем в любом интервале $I_{a,b}$, $0<a<b<\infty$, и $k\to\infty$. Суммируя эти соотношения по $x_{m,k}\in\in I_{a,b}$ и учитывая, что сумму справа можно записать как интегральную для функции $f(x)=(1/\sqrt{2\pi})x^{-3/2}\exp(-1/(2x))$ по отрезку $[a, b]$ с шагом разбиения $2(2k)^{-2}$ (исключая крайние отрезки разбиения), находим

$$\mathbf{P}(\omega: T_{2k}(\omega)/(2k)^2\in I_{a,b})\to\frac{1}{\sqrt{2\pi}}\int_a^b x^{-3/2}e^{-1/(2x)}\,dx.$$

Совершая в интеграле справа замену переменного $x=t^{-2}$, приведем его к виду

$$\frac{1}{\sqrt{2\pi}}\int_{1/\sqrt{b}}^{1/\sqrt{a}} t^3e^{-t^2/2}\,2t^{-3}\,dt=\sqrt{\frac{2}{\pi}}\int_{1/\sqrt{b}}^{1/\sqrt{a}} e^{-t^2/2}\,dt,$$

так что функция $f(x)=(1/\sqrt{2\pi})x^{-3/2}\exp(-1/(2x))$, $x>0$, $f(x)=0$, $x\leqslant 0$, — плотность вероятностного распределения. Элементарная выкладка показывает (см. 3.15), что

$$\mathbf{P}(\omega: T_{2k}(\omega)/(2k)^2\in I_{a,b})\to\int_a^b f(x)\,dx$$

для любых a, b. Для времен достижения нечетного уровня имеет место точно такое же предельное соотношение.

Полезно вспомнить, что сл. в. T_n, $n>0$, представляется как сумма n независимых случайных величин, распределенных так же, как $T_1: T_n=T_1+(T_2-T_1)+\ldots+(T_n-T_{n-1})$. Полагая $\tau_i=T_i-T_{i-1}$, $i=1,\ldots,n$, $T_0=0$, запишем полученный выше результат в виде

$$\mathbf{P}(\omega: (\tau_1(\omega)+\ldots+\tau_n(\omega))/n^2\in I_{a,b})\to\int_a^b f(x)\,dx.$$

Тот факт, что сумма независимых одинаково распределенных случайных величин растет примерно как квадрат числа слагаемых, заслуживает особого внимания. Закон больших чисел в форме Чебышева (см. 8.8) утверждает: когда независимые случайные величины имеют конечные и положительные математическое ожидание и дисперсию, их сумма растет примерно линейно. В рассматриваемом случае (см. 11.1)

$$\mathbf{P}(\omega : T_1(\omega) > 2n) \sim (\pi n)^{-1/2}, \quad n \to \infty,$$

так что математическое ожидание T_1 не существует (см. 7.3). Говорят также, что $\mathbf{M}T_1 = \infty$.

Нелинейный рост сумм $\tau_1 + \ldots + \tau_n$, $n = 1, 2, \ldots$, означает, что вклад в сумму различных слагаемых не равноценен. Выведем предельное распределение максимального слагаемого $\tau_{(n)} = \max\{\tau_k, k \leqslant n\}$. Запишем при $x > 0$ и $n \to \infty$

$$\mathbf{P}(\omega : \tau_{(n)}(\omega)/n^2 \leqslant x) = \mathbf{P}(\omega : T_1(\omega) \leqslant xn^2)^n \sim$$
$$\sim (1 - (\pi x n^2/2)^{-1/2})^n \to \exp(-(2/\pi)^{1/2} x^{-1/2}),$$

т. е. сл. в. $\tau_{(n)}/n^2$ имеет предельное распределение с ф. р. $F(x) = \exp(-(2/\pi)^{1/2} x^{-1/2})$, $x > 0$, $F(x) = 0$, $x \leqslant 0$. Следовательно, максимальный член суммы имеет порядок всей суммы. Можно показать, что сравнительно небольшое число старших порядковых статистик $\tau_{(n-k)}, \tau_{(n-k+1)}, \ldots, \tau_{(n)}$, n, $k \to \infty$, определяет всю сумму, так что вклад остальных слагаемых пренебрежимо мал.

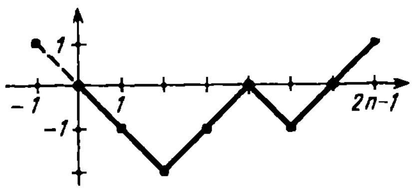

Рис. 28

11.8. Пример (предельная теорема для времен возвращения). Пусть $N_1, N_2, \ldots$ — последовательные моменты возвращения блуждающей частицы в начальное состояние. Из предыдущего вытекает (см. 11.3, 11.4), что сл. в. N_1, $N_2 - N_1$, $N_3 - N_2, \ldots$ независимы и одинаково распределены. Распределение сл. в. N_1 легко находится из принципа отражения 1.13, однако еще проще заметить, что любой путь длиной $2n - 1$, такой, что в момент $2n - 1$ происходит первое достижение положения 1 ($T_1 = 2n - 1$), можно преобразовать в путь длиной $2n$, добавляя к нему звено, соединяющее точку $(0, 0)$ с $(-1, 1)$ (рис. 28). Отсюда следует, что

$$2^{2n}\mathbf{P}(\omega : N_1(\omega) = 2n) = 2 \cdot 2^{2n-1}\mathbf{P}(\omega : T_1(\omega) = 2n - 1), \quad n = 1, 2, \ldots,$$

иначе говоря, распределение сл. в. N_1 совпадает с распределением сл. в. $T_1 + 1$. Но тогда распределение суммы $N_1 + (N_2 - N_1) + \ldots + (N_n - N_{n-1}) = N_n$ совпадает с распределением суммы $T_1 + 1 + (T_2 - T_1 + 1) + \ldots + (T_n - T_{n-1} + 1) = T_n + n$, т. е. сл. в. N_n и $T_n + n$ одинаково распределены.

Из 11.7 получаем

$$\lim_{n \to \infty} \mathbf{P}(\omega : N_n(\omega)/n^2 \leqslant x) = \lim_{n \to \infty} \mathbf{P}(\omega : T_n(\omega)/n^2 \leqslant x - 1/n^2).$$

Переходя в неравенствах

$$F_{T_n/n^2}(x-\varepsilon) \leqslant F_{T_n/n^2}(x-1/n^2) \leqslant F_{T_n/n^2}(x), \quad 1/n^2 < \varepsilon,$$

к пределу сначала по $n \to \infty$, а затем по $\varepsilon \to 0$ и используя непрерывность ф. р. $F(x) = \lim\limits_{n\to\infty} F_{T_n/n^2}(x)$, получаем

$$\lim_{n\to\infty} F_{T_n/n^2}(x-1/n^2) = F(x),$$

следовательно,

$$\lim_{n\to\infty} F_{N_n/n^2}(x) = F(x).$$

Из предельной теоремы для времен возвращения N_n выводится предельная теорема для числа возвращений ν_n блуждающей частицы в начальное состояние в течение первых n шагов:

$$\nu_n(\omega) = m \Leftrightarrow N_m(\omega) \leqslant n, \quad N_{m+1}(\omega) > n, \quad n, m = 1, 2, \ldots .$$

Соотношение, связывающее последовательности ν_n, $n=1, 2, \ldots$, и N_n, $n=1, 2, \ldots$, удобнее записать для этих целей в виде

$$\nu_n(\omega) \geqslant m \Leftrightarrow N_m(\omega) \leqslant n, \quad n, m = 1, 2, \ldots .$$

Полагая $n/m^2 = x_{n,m}$, находим отсюда

$$\mathbf{P}(\omega : N_m(\omega)/m^2 \leqslant x_{n,m}) = \mathbf{P}\left(\omega : \nu_n(\omega)/\sqrt{n} \geqslant x_{n,m}^{-1/2}\right).$$

Пусть $n \to \infty$ и выберем $m = m_n$ так, чтобы $x_{n,m} \to x$, где x — любое фиксированное. В таком случае

$$\lim_{n\to\infty} \mathbf{P}\left(\omega : \nu_n(\omega)/\sqrt{n} \geqslant x_{n,m}^{-1/2}\right) = F(x),$$

или

$$\lim_{n\to\infty} F_{\nu_n/\sqrt{n}}(x_{n,m}) = 1 - F(x^{-2}).$$

Записав это соотношение для точек $x \pm \varepsilon$ и учитывая монотонность функции распределения, получаем

$$1 - F\left((x-\varepsilon)^{-2}\right) \leqslant \varliminf_{n\to\infty} F_{\nu_n/\sqrt{n}}(x) \leqslant \varlimsup_{n\to\infty} F_{\nu_n/\sqrt{n}}(x) \leqslant 1 - F\left((x+\varepsilon)^{-2}\right).$$

Ввиду произвольности $\varepsilon > 0$ и непрерывности $F(x)$ окончательно получаем

$$\lim_{n\to\infty} F_{\nu_n/\sqrt{n}}(x) = 1 - F(x^{-2}).$$

Так как (см. 11.7)

$$F(b) - F(a) = \sqrt{\frac{2}{\pi}} \int\limits_{1/\sqrt{b}}^{1/\sqrt{a}} e^{-t^2/2}\, dt,$$

то

$$1-F(x^{-2})=\sqrt{\frac{2}{\pi}}\int_0^x e^{-t^2/2}\,dt,$$

т. е. предельное распределение для $\nu_n/\sqrt{n}$ есть распределение модуля стандартной нормальной величины.

Следует обратить внимание на то, что число возвращений за время n растет примерно как $\sqrt{n}$, а вовсе не пропорционально n, как можно было бы думать. Из предыдущего ясно, что это связано с тем, что относительно небольшое число очень длинных (порядка n) времен возвращения будет составлять подавляющую часть времени блуждания n. Отметим, кстати, что время $n-N_{\nu_n}$ от последнего перед n момента возвращения до n имеет порядок n, так как (см. 3.16 и далее)

$$F_{N(\nu_n)/n}(x)\to 2\pi^{-1}\arcsin\sqrt{x},\quad 0<x<1\quad (N(\nu_n)\equiv N_{\nu_n}).$$

11.9. Пример (ветвящийся процесс). Рассмотрим отрезок случайного блуждания S_n, $n=1, 2, \ldots, S_0=0$, вплоть до момента T_{-1} первого достижения точки -1. С отрезком S_n, $n\leqslant T_{-1}$, свяжем последовательность сл. в. Z_k, $k=0, 1, 2, \ldots$, полагая Z_k равным числу переходов блуждающей частицы из положения k в $k-1$:

$$Z_k(\omega)=\sum_{\{n\leqslant T_{-1}(\omega)\}} I_{\{\omega: S_{n-1}(\omega)=k,\, S_n(\omega)=k-1\}}(\omega).$$

Покажем, что последовательность $Z_0=1$, Z_1, $Z_2, \ldots$ образует ветвящийся процесс Гальтона—Ватсона (см. 9.24).

Проведем сначала неформальное рассуждение. Пусть $N_i(\omega)$, $i=1, 2, \ldots,$ — последовательные моменты возвращения в начальное состояние. Дополним траекторию блуждания S_n, $0\leqslant n\leqslant T_{-1}$, еще одним звеном, соединяющим точку $(0, 0)$ с точкой $(-1, -1)$. Назовем начальной «частицей» в процессе размножения отрезок с концами $(-1, -1)$, $(T_{-1}, -1)$ (рис. 29). Непосредственными потоками начальной частицы назовем отрезки на оси абсцисс с концами $N_{i-1}(\omega)$, $N_i(\omega)$, $i=1, 2, \ldots, Z_1$ $(N_0(\omega)=0)$. Чтобы определить непосредственных потомков i-й частицы, переместим начало координат в точку $(N_{i-1}+1, 1)$ и рассмотрим отрезок блуждания от $N_{i-1}(\omega)$ до $N_i(\omega)$. Из сравнения рис. 29 и 30 видно, что на рис. 30 повторяется начальная ситуация. Назовем непосредственными по-

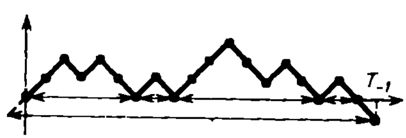

Рис. 29

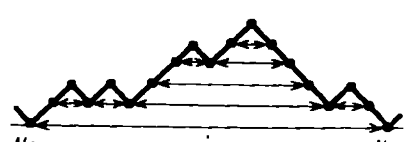

Рис. 30

томками i-й частицы соответствующие отрезки на новой оси абсцисс. Продолжая этот процесс, построим систему горизонтальных отрезков, находящихся во взаимно однозначном соответствии с переходами блуждающей частицы из какого-либо положения k в $k-1$. Суммарное число отрезков, лежащих на горизонтали $y=k-1$, совпадает с Z_k, $k=0, 1, 2, \ldots$. Теперь заметим, что отрезки блуждания $S_j(\omega)$, $N_{i-1}(\omega) \leqslant j \leqslant N_i(\omega)$, $i=1, 2, \ldots$, независимы и одинаково распределены (см. 11.3, 11.4). Поэтому все процессы размножения каждой из Z_1 частиц первого поколения не зависят друг от друга, а закон размножения от поколения к поколению не меняется.

Нетрудно понять, что

$$\mathbf{P}(\omega : Z_1(\omega)=m)=2^{-m-1}, \quad m=0, 1, 2, \ldots .$$

В самом деле,

$$\mathbf{P}(\omega : Z_1(\omega)=0)=\mathbf{P}(\omega : S_1(\omega)=-1)=2^{-1}, \ \mathbf{P}(\omega : Z_1(\omega)=1)=$$

$$=\mathbf{P}(\omega : S_1(\omega)=1, \ S_{N_1+1}(\omega)=-1)=2^{-2}, \ \mathbf{P}(\omega : Z_1(\omega)=2)=$$

$$=\mathbf{P}(\omega : S_1(\omega)=1, \ S_{N_1+1}(\omega)=1, \ S_{N_2+1}(\omega)=-1)=2^{-3}, \ \ldots$$

Наконец, заметим, что переходные вероятности марковской цепи $Z_0, Z_1, Z_2, \ldots$ равны

$$p_{Z_2|Z_1}(j|i)=C^{j}_{i+j-1}2^{-i-j}, \quad j=0, 1, 2, \ldots, \ i>0,$$

так как сумма i независимых геометрически распределенных случайных величин имеет отрицательное биномиальное распределение (см. 4.6).

Формализация проведенных выше рассуждений довольно громоздка, значительно проще выглядит непосредственный подсчет числа путей длины $n=2(i_1+i_2+\ldots+i_k)+1$, удовлетворяющих условию $Z_0=i_1$, $Z_2=i_2, \ldots, Z_k=i_k$, $Z_{k+1}=0$, $i_1>0$, $i_2>0, \ldots, i_k>0$. Обратимся к построению на рис. 29, 30, где отрезку пути S_n, $n \leqslant T_{-1}$, сопоставлена система горизонтальных отрезков на координатной плоскости. Нетрудно понять, что по этой системе путь восстанавливается однозначно. Более того, для однозначного восстановления пути знание абсцисс концов (a, i), (b, i) любого из отрезков системы не обязательно. Достаточно, чтобы для каждого такого отрезка было указано количество отрезков $[(c, i+1), (d, i+1)]$, лежащих непосредственно над ним: $a \leqslant c < d \leqslant b$. В терминах описанного выше процесса размножения «частиц» это означает, что для каждой частицы должно быть указано число ее непосредственных потомков.

Для подсчета путей с заданными значениями $Z_i>0$, $i=1, \ldots, k$, $Z_{k+1}=0$ применима комбинаторная схема разложения неразличимых шаров по ящикам. На первом шаге подсчета объявим «ящиками» $Z_1=i_1$ отрезков, лежащих на оси абсцисс, а «шарами» — $Z_2=i_2$ отрезков, лежащих на горизонтальной прямой $y=1$. Число различных разложений (т. е. расположений Z_2 отрезков второго яруса так, что каждый из них проектируется на некоторый из Z_1 отрезков первого яруса), равно (см. 1.26) $C_{i_1+i_2-1}^{i_2}$. Точно так же получаем $C_{i_2+i_3-1}^{i_3}$ расположений $Z_3=i_3$ отрезков третьего яруса над $Z_2=i_2$ отрезками второго и т. д. В результате находим число путей:

$$C_{i_1+i_2-1}^{i_2} C_{i_2+i_3-1}^{i_3} \ldots C_{i_{k-1}+i_k-1}^{i_k}.$$

Соответствующая вероятность получается домножением на 2^{-n}, $n=2(i_1+i_2+\ldots+i_k)+1$. Запишем ее в виде

$$\mathbf{P}(\omega: Z_{i_1}(\omega)=i_1>0, \ldots, Z_k(\omega)=i_k>0, Z_{k+1}(\omega)=0)=$$
$$=p(i_1|1)\,p(i_2|i_1)\ldots p(i_k|i_{k-1})\,p(0|i_k), \quad p(j|i)=C_{i+j-1}^{j}2^{-i-j}.$$

Чтобы убедиться, что последовательность сл. в. $Z_0, Z_1, Z_2, \ldots$ образует марковскую цепь, остается проверить, что

$$\mathbf{P}(\omega: Z_1(\omega)=i_1>0, \ldots, Z_k(\omega)=i_k>0)=p(i_1|1)\ldots p(i_k|i_{k-1}).$$

Так как сл. в. $T_{-1}(\omega)$ конечна с вероятностью 1, то с вероятностью 1 последовательность $Z_k(\omega)$, $k=0, 1, 2, \ldots$, начиная с некоторого номера (своего для каждого ω), обращается в нуль, т. е.

$$1=\sum_{k=0}^{\infty}\sum_{(i_1>0,\ldots,i_k>0)}\mathbf{P}(\omega: Z_1(\omega)=i_1, \ldots, Z_k(\omega)=i_k, Z_{k+1}(\omega)=0)=$$
$$=\sum_{k=0}^{\infty}\sum_{(i_1>0,\ldots,i_k>0)} p(i_1|1)\,p(i_2|i_1)\ldots p(i_k|i_{k-1})\,p(0|i_k)=$$
$$=p(0|1)+\sum_{i_1=1}^{\infty}p(i_1|1)\sum_{k=1}^{\infty}\sum_{(i_2>0,\ldots,i_k>0)} p(i_2|i_1)\ldots p(0|i_k).$$

Поскольку

$$p(0|1)+\sum_{i=1}^{\infty}p(i|1)=1, \quad p(i|1)>0, \quad i=1, 2, \ldots,$$

выводим отсюда, что при любом $i_1>0$

$$\sum_{k=1}^{\infty}\sum_{(i_2>0,\ldots,i_k>0)} p(i_2|i_1)\ldots p(i_k|i_{k-1})\,p(0|i_k)=1.$$

Поэтому, разбивая событие $\{\omega: Z_1(\omega)=i_1, \ldots, Z_k(\omega)=i_k\}$ на части

$$\{\omega: Z_1(\omega)=i_1, Z_2(\omega)=i_2, \ldots, Z_{k+l}(\omega)=i_{k+l}>0, Z_{k+l+1}(\omega)=0\},$$

получаем

$$\mathbf{P}(\omega : Z_1(\omega)=i_1, \ldots, Z_k(\omega)=i_k>0)=p(i_1|1)\ldots p(i_k|i_{k-1})\times$$

$$\times\sum_{l=0}^{\infty}\sum_{\{i_{k+1}>0,\ldots,i_{k+l}>0\}} p(i_{k+1}|i_k)\ldots p(i_{k+l}|i_{k+l-1})\,p(0|i_{k+l})=$$

$$=p(i_1|1)\ldots p(i_k|i_{k-1}).$$

Итак, последовательность сл. в. $Z_0=1$, Z_1, $Z_2, \ldots$ образует марковскую цепь с переходными вероятностями

$$p(j|i)=C^j_{i+j-1}\,2^{-i-j}, \quad i>0, \; j=0, 1, 2, \ldots, \; p(0|0)=1.$$

Поскольку отрицательное биномиальное распределение $p(j|i)$, $i>1$, — распределение суммы i независимых, распределенных по геометрическому закону $p(j|1)$, случайных величин, то условное распределение Z_{k+1} при условии $Z_k=i$ совпадает с распределением суммы i независимых случайных величин, распределенных также, как Z_1. Тем самым установлено, что Z_k, $k=0, 1, 2, \ldots$, — ветвящийся процесс.

Ветвящийся процесс $Z_0=1$, Z_1, $Z_2, \ldots$ — *вырождающийся*: с вероятностью 1, начиная с некоторого случайного момента времени, все $Z_k=0$. Поэтому с вероятностью 1 конечна сл. в. $U=Z_0+Z_1+$ $+Z_2+\ldots$ — суммарное число когда-либо живших частиц. Интерес к сл. в. U связан с некоторыми физическими и биологическими процессами, в которых «размножающиеся» частицы наряду с непосредственными потомками — частицами такого же вида — производят одну так называемую *финальную* частицу, которая в дальнейшем остается неизменной и не размножается. В таком случае U представляет собой суммарное количество «вещества», являющегося продуктом ветвящегося процесса.

Распределение вероятностей сл. в. U находится из равенства $2U-1=T_{-1}$. Представим себе, что ветвящийся процесс начался с $Z_0=n$ частиц, каждая из которых размножается независимо. Суммарное число $U^{(n)}$ всех живших частиц в этом процессе равно $U^{(n)}=U_1+U_2+\ldots+U_n$, где U_i, $i=1, 2, \ldots, n$, независимы и распределены так же, как U. Следовательно, $U^{(n)}$ распределена как $(T_n+n)/2$, где T_n — время достижения симметричным простым блужданием $S_0=0$, S_1, $S_2, \ldots$ положения n. Таким образом, из 11.8 получаем

$$\mathbf{P}(\omega : U^{(n)}/n^2\leqslant x)\to F(2x), \quad n\to\infty,$$

где ф. р. $F(x)$ имеет плотность $f(x)=\left(1/\sqrt{2\pi}\right)x^{-3/2}e^{-1/(2x)}$, $x>0$. Обращаясь к модели процесса с финальными частицами, приходим к интересному заключению: количество конечного «продукта» примерно пропорционально квадрату начального количества вещества.

11.10. Пример (условное случайное блуждание и броуновский мост). Пусть $(\Omega, \mathscr{A}, \mathbf{P})$ — вероятностное пространство, на котором определено простое симметричное блуждание S_k, $k=0, 1, 2, \ldots$. Рассмотрим какое-либо событие Λ_n, связанное с отрезком блуждания S_k, $k \leqslant n$, и образуем вероятностное пространство $(\Lambda_n, \mathcal{A} \cap \Lambda_n, \mathbf{P}_{\Lambda_n})$, где $\mathcal{A} \cap \Lambda_n$ есть класс событий вида $A \cap \Lambda_n$, $A \in \mathcal{A}$, $\mathbf{P}_{\Lambda_n}$ — условная вероятностная мера при условии события Λ_n. Последовательность $S_k(\omega)$, $k \leqslant n$, рассматриваемую как последовательность случайных величин на $(\Lambda_n, \mathcal{A} \cap \Lambda_n, \mathbf{P}_{\Lambda_n})$, называют *условным случайным блужданием* при условии события Λ_n и обозначают $(S_k | \Lambda_n)(\omega)$, $k \leqslant n$.

Исследуем условное случайное блуждание при условии $\Lambda_n = \{\omega : S_n(\omega) = 0\}$, предполагая n четным. Полагая для краткости $Y_k = (S_k | \Lambda_n)$, запишем при $m \leqslant n$, $|i_{k+1} - i_k| = 1$, $k \leqslant m$, $i_0 = i_n = 0$:

$$p_{Y_1, \ldots, Y_m}(i_1, \ldots, i_m) = \mathbf{P}(\omega : S_k(\omega) = i_k,\ k \leqslant m,\ S_n(\omega) = 0) \times$$

$$\times \mathbf{P}(\omega : S_n(\omega) = 0)^{-1} = \mathbf{P}(\omega : S_k(\omega) = i_k,\ k \leqslant m)\, \mathbf{P}(\omega : S_n(\omega) - S_m(\omega) =$$

$$= -i_m)\, \mathbf{P}(\omega : S_n(\omega) = 0)^{-1} = 2^{-m} \cdot N_{n-m, i_m} \cdot 2^{-n+m} \cdot N_{n,0}^{-1} 2^n =$$

$$= N_{n-m, i_m} / N_{n,0}, \quad N_{n,k} = C_n^{(n+k)/2},$$

откуда получаем

$$p_{Y_{m+1} | Y_1, \ldots, Y_m}(i_{m+1} | i_1, \ldots, i_m) = \mathbf{P}(\omega : Y_k(\omega) = i_k,\ k \leqslant m+1) \times$$

$$\times \mathbf{P}(\omega : Y_k(\omega) = i_k,\ k \leqslant m)^{-1} = N_{n-m-1, i_{m+1}} / N_{n-m, i_m}.$$

Таким образом, последовательность сл. в Y_k, $k \leqslant n$, образует марковскую цепь с переходными вероятностями

$$p_{Y_{m+1} | Y_m}(i \pm 1 | i) \equiv p_m(i \pm 1 | i),$$

зависящими от номера шага m. Такие марковские цепи называются *неоднородными*. Легко подсчитать, что

$$p_m(i \pm 1 | i) = 2^{-1}(1 \mp i/(n-m)), \quad |i| \leqslant \min(m, n-m).$$

В область $|i| > \min(m, n-m)$ частица на m-м шаге не попадает. Интересно проследить, как условие $S_n = 0$ сказывается на движении блуждающей частицы. Если в случайном блуждании частица с равными вероятностями перемещается из положения i в $i \pm 1$, то теперь вероятности перемещений зависят от положения i и момента времени (шага блуждания) m. При $i > 0$ $p_m(i-1 | i) > p_m(i+1 | i)$, а при $i < 0$ $p_m(i+1 | i) > p_m(i-1 | i)$, т. е. частица как бы притягивается к начальному положению, причем степень притяжения — асимметрия вероятностей сдвига — возрастает с увеличением отклонения $|i|$ при данном m.

Со случайным блужданием S_k, $k=0, 1, 2, \ldots$, в разделах 3.19, 4.27 была связана последовательность случайных процессов с непрерывным временем

$$W_t^{(n)}=S_{[nt]}/\sqrt{n}, \quad t\geqslant 0,$$

где $[a]$ обозначает целую часть числа a. Было показано, что

$$\mathbf{P}(\omega : W_{t_i}^{(n)}(\omega)\leqslant x_i, \ i=1, \ldots, k)\to F_{t_1,\ldots,t_k}(x_1, \ldots, x_k), \ n\to\infty,$$

при любых $0<t_1<\ldots<t_k$, $x_1,\ldots,x_k$, $k=1, 2, \ldots$, где ф. р. $F_{t_1,\ldots,t_k}$ имеет плотность

$$f_{t_1,\ldots,t_k}(x_1, \ldots, x_k)=\prod_{i=1}^{k}\left(1/\sqrt{t_i-t_{i-1}}\right)\varphi\left((x_i-x_{i-1})/\sqrt{t_i-t_{i-1}}\right),$$

$$x_0=t_0=0,$$

$\varphi(x)$ — стандартная нормальная плотность. Случайный процесс W_t, $t\geqslant 0$, конечномерные распределения которого задаются плотностями $f_{t_1,\ldots,t_k}$, называется *стандартным броуновским движением* (доказательство существования такого процесса проведено в 10.41). Говорят, что последовательность процессов $W_t^{(n)}$, $t\geqslant 0$, *сходится по распределению* к процессу W_t, $t\geqslant 0$.

Рассмотрим по аналогии процесс

$$(W_t^{(n)}\,|\,\Lambda_n)=\left(S_{[nt]}/\sqrt{n}\,|\,\Lambda_n\right).$$

Хотелось бы заключить, что последовательность процессов $(W_t^{(n)}|\Lambda_n)$, $0\leqslant t\leqslant 1$, сходится по распределению к процессу W_t, $0\leqslant t\leqslant 1$, подчиненному условию $W_1=0$. Условный процесс $W_t^0=(W_t|\{W_1=0\})$, мы определим как процесс, конечномерные плотности которого являются условными плотностями процесса броуновского движения:

$$f_{W^0(t_1),\ldots,W^0(t_k)}(x_1, \ldots, x_k)=f_{t_1,\ldots,t_k,1}(x_1, \ldots, x_k, 0)\cdot\sqrt{2\pi},$$

$$0\leqslant t_1<\ldots<t_k\leqslant 1.$$

Покажем, что действительно такая сходимость имеет место. Пусть $n_0=0<n_1<n_2<\ldots<n_k\leqslant n=n_{k+1}$. Действуя аналогично предыдущему, запишем

$$\mathbf{P}(\omega : Y_{n_i}(\omega)=m_i, \ i\leqslant k)=\mathbf{P}(\omega : S_{n_i}(\omega)=m_i, \ i\leqslant k, \ S_n(\omega)=0)\times$$

$$\times\mathbf{P}(\omega : S_n(\omega)=0)^{-1}=\prod_{i=1}^{k+1}N_{n_i-n_{i-1},m_i-m_{i-1}}\cdot N_{n,0}^{-1}, \quad m_0=m_{k+1}=0.$$

Поскольку (см. 3.19)

$$N_{[nt],[x\sqrt{n}]}2^{-[nt]}\sim\left(2/\sqrt{nt}\right)\varphi\left(x/\sqrt{t}\right), \ n\to\infty,$$

то при $n_i=[nt_i]$, $m_i=[x_i\sqrt{n}]$, $i=1,\ldots,k$, $t_0=0<t_1<\ldots<t_k<1=t_{k+1}$, и любых $x_1,\ldots,x_k$, получаем (полагая $x_{k+1}=0$):

$$\mathbf{P}(\omega:Y_{n_i}(\omega)=m_i,\ i\leqslant k)\sim\prod_{i=1}^{k+1}\frac{2}{(n(t_i-t_{i-1}))^{1/2}}\,\varphi\left(\frac{x_i-x_{i-1}}{(t_i-t_{i-1})^{1/2}}\right)\times$$

$$\times(2/(n\pi))^{-1}=(2/\sqrt{n})^k f_{t_1,\ldots,t_k,1}(x_1,\ldots,x_k,0)\cdot\sqrt{2\pi}.$$

Стандартные рассуждения приводят к интегральной предельной теореме

$$\mathbf{P}\left(\omega:Y_{n_i}(\omega)/\sqrt{n}\in I_{a_i,b_i},\ i\leqslant k\right)\to p_{W^0(t_1),\ldots,W^0(t_k)}(I_{a_1,b_1},\ldots,I_{a_k,b_k}).$$

Структура предельного процесса $W^0(t)$, $0\leqslant t\leqslant 1$, становится более понятной после некоторого преобразования его конечномерных плотностей. Рассмотрим условную плотность

$$f_{W^0(t_{k+1})|W^0(t_1),\ldots,W^0(t_k)}(x_{k+1}|x_1,\ldots,x_k)=$$

$$=f_{t_1,\ldots,t_{k+1},1}(x_1,\ldots,x_{k+1},0)/f_{t_1,\ldots,t_k,1}(x_1,\ldots,x_k,0)=$$

$$=\frac{1}{(t_{k+1}-t_k)^{1/2}}\,\varphi\left(\frac{x_{k+1}-x_k}{(t_{k+1}-t_k)^{1/2}}\right)\left(\frac{1-t_k}{1-t_{k+1}}\right)^{1/2}\cdot\frac{\varphi(-x_{k+1}/\sqrt{1-t_{k+1}})}{\varphi(-x_k/\sqrt{1-t_k})}.$$

Так как она не зависит от $x_1,\ldots,x_{k-1}$, то последовательность сл. в. $W^0(t_1)$, $W^0(t_2),\ldots$ при любых $0<t_1<t_2<\ldots$ образует марковскую цепь. Ее переходная плотность равна

$$f_{W^0(t)|W^0(s)}(y|x)=\left(\frac{1-s}{(t-s)(1-t)}\right)^{1/2}\frac{\varphi(y/\sqrt{1-t})}{\varphi(x/\sqrt{1-s})}\,\varphi\left(\frac{y-x}{\sqrt{t-s}}\right),$$

где $0<s<t<1$ и мы воспользовались четностью функции $\varphi(x)$. Обозначим правую часть $p_0(s,x,t,y)$. Так как

$$f_{W^0(t)}(x)=f_{t,1}(x,0)=(1/\sqrt{t})\,\varphi(x/\sqrt{t})\,(1/\sqrt{1-t})\,\varphi(-x/\sqrt{1-t})=$$

$$=\frac{1}{2\pi(t(1-t))^{1/2}}\cdot\exp\left(-\frac{1}{2}\left(\frac{x^2}{t}-\frac{x^2}{1-t}\right)\right)=$$

$$=(1/\sqrt{2\pi t(1-t)})\,\varphi(x/\sqrt{t(1-t)})=(1/\sqrt{2\pi})p_0(0,0,t,x),$$

то конечномерную плотность процесса $W^0(t)$ можно записать в виде

$$f_{W^0(t_1),\ldots,W^0(t_k)}(x_1,\ldots,x_k)=\prod_{i=0}^{k-1}p_0(t_i,x_i,t_{i+1},x_{i+1}),$$

где $t_0=0<t_1<\ldots<t_k<1$, $x_0=0$.

Недостающим звеном в проведенных выше рассуждениях является вопрос о существовании процесса $W^0(t)$ (см. 10.41). В следующей задаче будет показано, что $W^0(t)$ может быть просто выражен через $W(t)$. В пределах данного примера необходимо подчеркнуть, что формула для $f_{W^0(t_1),\ldots,W^0(t_k)}$ в самом деле задает k-мерное вероятностное распределение, т. е. интеграл от f по R^k равен единице. Это непосредственно следует из того, что функция $p_0(s, x, t, y)$ — одномерная плотность по y:

$$\int_{-\infty}^{\infty} p_0(s, x, t, y)\,dy = \sqrt{1-s}\cdot\varphi\left(x/\sqrt{1-s}\right)^{-1}\times$$

$$\times\int_{-\infty}^{\infty}\left(1/\sqrt{t-s}\right)\varphi\left((y-x)/\sqrt{t-s}\right)\left(1/\sqrt{1-t}\right)\varphi\left(y/\sqrt{1-t}\right)dy=1,$$

так как интеграл представляет собой свертку двух нормальных плотностей.

Процесс W_t^0, $0\leqslant t\leqslant 1$, называется *броуновским мостом*. Свойство независимости условных плотностей

$$f_{W^0(t_{k+1})|W^0(t_1),\ldots,W^0(t_k)}\,(x_{k+1}|x_1, \ldots, x_k)$$

от переменных $x_1,\ldots,x_{k-1}$ при любых $0<t_1<\ldots<t_{k+1}\leqslant 1$, $k=1, 2,\ldots$, называется *марковским свойством* процесса W_t^0, $0\leqslant t\leqslant 1$, а любой процесс, обладающий этим свойством, называется *марковским*. Функция

$$p_0(s, x, t, y)=f_{W^0(t)|W^0(s)}\,(y|x)$$

называется *переходной плотностью* марковского процесса W_t^0, $0\leqslant t\leqslant 1$. Формула для конечномерных плотностей

$$f_{W^0(t_1),\ldots,W^0(t_k)}\,(x_1, \ldots, x_k)=f_{W^0(t_1)}(x_1)\prod_{i=1}^{k-1}p_0(t_i, x_i, t_{i+1}, x_{i+1})$$

следует из марковского свойства процесса, которое в свою очередь вытекает из этой формулы.

Броуновское движение также, очевидно, является марковским процессом. Его переходная плотность

$$p(s, x, t, y)=\left(1/\sqrt{t-s}\right)\varphi\left((y-x)/\sqrt{t-s}\right),\quad 0\leqslant s<t,$$

зависит от t и s через их разность $t-s$ — это *однородный марковский процесс*. Тот факт, что $p(s, x, t, y)$ зависит от x и y через их разность $y-x$, есть свойство независимости приращений процесса $W(t)$.

11.11.* Задача. Пусть $W(t)$ — стандартное броуновское движение, $W^0(t)$, $0\leqslant t\leqslant 1$, — броуновский мост. Показать, что $W(t)-tW(1)$, $0\leqslant t\leqslant 1$ и $W^0(t)$, $0\leqslant t\leqslant 1$, имеют одинаковые конечномерные распределения, т. е., как говорят, *совпадают по распределению*.

[Пусть $0<t_1<\ldots<t_k<1$ и положим

$$Y_i=W(t_i)-t_iW(1),\ i=1,\ldots,k,\ Y_{k+1}=W(1).$$

По формуле 3.26 преобразования плотности при линейном преобразовании случайных величин имеем

$$f_{Y_1,\ldots,Y_{k+1}}(y_1,\ \ldots,\ y_{k+1})=f_{W(t_1),\ldots,W(t_k),W(1)}(x_1,\ \ldots,\ x_{k+1}),$$

где переменные $y_1,\ldots,y_{k+1}$ и $x_1,\ldots,x_{k+1}$ связаны линейным преобразованием с единичным детерминантом:

$$y_i=x_i-t_ix_{k+1},\quad i=1,\ldots,k,\quad y_{k+1}=x_{k+1},$$

или

$$x_i=y_i+t_iy_{k+1},\quad i=1,\ldots,k,\quad x_{k+1}=y_{k+1}.$$

Тем самым для плотности $f_{Y_1,\ldots,Y_{k+1}}$ получаем выражение

$$\prod_{i=1}^{k}\frac{1}{\sqrt{t_i-t_{i-1}}}\varphi\left(\frac{y_i-y_{i-1}+(t_i-t_{i-1})y_{k+1}}{\sqrt{t_i-t_{i-1}}}\right)\times$$

$$\times\frac{1}{\sqrt{1-t_k}}\varphi\left(\frac{y_{k+1}(1-t_k)-y_k}{\sqrt{1-t_k}}\right)=\left(\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi(t_i-t_{i-1})}}\right)\cdot\frac{1}{\sqrt{2\pi(1-t_k)}}\times$$

$$\times\exp\left\{-\frac{1}{2}\left(\sum_{i=1}^{k}(\sqrt{t_i-t_{i-1}}\,y_{k+1}+(y_i-y_{i-1})/\sqrt{t_i-t_{i-1}})^2+\right.\right.$$

$$\left.\left.+(\sqrt{1-t_k}\,y_{k+1}-y_k/\sqrt{1-t_k})^2\right)\right\},$$

где мы полагаем $t_0=y_0=0$. Преобразуем сумму под знаком экспоненты к виду

$$\sum_{i=1}^{k}(t_i-t_{i-1})y_{k+1}^2+2\sum_{i=1}^{k}y_{k+1}(y_i-y_{i-1})+\sum_{i=1}^{k}\frac{(y_i-y_{i-1})^2}{t_i-t_{i-1}}+$$

$$+(1-t_k)y_{k+1}^2-2y_{k+1}y_k+\frac{y_k^2}{1-t_k}=y_{k+1}^2+\sum_{i=1}^{k}\frac{(y_i-y_{i-1})^2}{t_i-t_{i-1}}+\frac{y_k^2}{1-t_k}.$$

Плотность $f_{Y_1,\ldots,Y_k}$ получим, интегрируя $f_{Y_1,\ldots,Y_{k+1}}(y_1,\ \ldots,\ y_{k+1})$ по переменному y_{k+1}. С учетом проведенного преобразования приходим к следующему выражению для $f_{Y_1,\ldots,Y_k}(y_1,\ \ldots,\ y_k)$:

$$\left(\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi(t_i-t_{i-1})}}\right)\cdot\frac{1}{\sqrt{1-t_k}}\exp\left\{-\frac{1}{2}\left(\sum_{i=1}^{k}\frac{(y_i-y_{i-1})^2}{t_i-t_{i-1}}+\frac{y_k^2}{1-t_k}\right)\right\}.$$

которое, очевидно, совпадает с

$$\sqrt{2\pi}\cdot f_{W(t_1),\ldots,W(t_k),W(1)}(x_1,\ \ldots,\ x_k,\ 0)=f_{W^0(t_1),\ldots,W^0(t_k)}(x_1,\ \ldots,\ x_k).]$$

11.12. Задача. Положим $M_n=\max\{S_k,\ k\leqslant n\}$, $\Lambda_n=\{\omega: S_n(\omega)=0\}$. Вывести из 1.17 предельную теорему для $(M_n|\Lambda_n)=$ $=\max\{(S_k|\Lambda_n),\ k\leqslant n\}$, n — четное.

$$[\mathbf{P}(\omega: M_n(\omega)\geqslant r,\ S_n(\omega)=0)/\mathbf{P}(\omega: S_n(\omega)=0)=N_{n,\,2r}/N_{n,\,0}.$$

Полагая $r=[x\sqrt{n}]$), получаем с учетом 3.19

$$\mathbf{P}(\omega: n^{-1/2}(M_n|\Lambda_n)(\omega)\leqslant x)\to 1-e^{-2x^2},\ x\geqslant 0.]$$

11.13.* Задача. Пусть $\tau_n(\omega)$ — наименьший момент времени, такой, что $S_{\tau_n}(\omega)=S_{\tau_n+n}(\omega)=0$. Показать, что отрезок S_k, $\tau_n\leqslant$ $\leqslant k\leqslant\tau_n+n$, случайного блуждания имеет такое же распределение вероятностей, как условное случайное блуждание $(S_k|\Lambda_n)$, $k\leqslant n$, $\Lambda_n=\{\omega: S_n(\omega)=0\}$ (n — четно).

[Заметим прежде всего, что $\mathbf{P}(\omega:\tau_n(\omega)<\infty)=1$. В самом деле, пусть T_k, $k=1,\ 2,\ldots,$ — последовательные моменты возвращения случайного блуждания S_k, $k=0,\ 1,\ 2,\ldots,$ в нуль, $T_0=S_0=0$. Тогда сл. в. $\Delta_k=T_k-T_{k-1}$, $k=1,\ 2,\ldots,$ независимы и одинаково распределены, причем $\mathbf{P}(\omega:\Delta_1(\omega)=n)\equiv p_n>0$ при любом четном $n>0$. Поэтому

$$\mathbf{P}(\omega: \tau_n(\omega)<\infty)=\mathbf{P}\Big(\bigcup_{k=1}^{\infty}\{\omega: \Delta_i(\omega)\neq n,\ i<k,\ \Delta_k(\omega)=n\}\Big)=$$

$$=\sum_{k=1}^{\infty}(1-p_n)^{k-1}p_n=1.$$

Пусть $|j_k-j_{k-1}|=1$, $0\leqslant k\leqslant n$, $j_0=j_n=0$. Разложив событие $\{\omega: S_{\tau_n+k}(\omega)=j_k,\ 0\leqslant k\leqslant n\}$ по полной группе событий $\{\omega:\tau_n(\omega)=$ $=T_i(\omega)\}$, $i=1,\ 2,\ \ldots$ (плюс событие $\tau_n(\omega)=\infty$, имеющее меру нуль), получаем для его вероятности выражение

$$\sum_{i=1}^{\infty}\mathbf{P}(\omega: S_{\tau_n+k}(\omega)=j_k,\ 0\leqslant k\leqslant n,\ \tau_n(\omega)=T_i(\omega)=$$

$$=\sum_{i=1}^{\infty}\mathbf{P}(\omega: S_{T_i+k}(\omega)=j_k,\ 0\leqslant k\leqslant n,\ \Delta_l(\omega)\neq n,\ l\leqslant i),$$

где мы полагаем $\Delta_0(\omega)=0$, так что $\{\omega:\Delta_0(\omega)\neq n\}=\Omega$, $n>0$. Сл. в. $\Delta_1,\ldots,\Delta_{i-1}$ связаны с поведением случайного блуждания до момента T_i и потому независимы от S_{T_i+k}, $k=0,\ 1,\ 2,\ \ldots$ (см. 11.5), к тому же S_{T_i+k}, $k=0,\ 1,\ 2,\ \ldots,$ представляет собой простое симметричное блуждание, так что

$$\mathbf{P}(\omega : S_{\tau_n+k}(\omega)=j_k,\ 0\leqslant k\leqslant n)=\mathbf{P}(\omega : S_k=j_k,\ 0\leqslant k\leqslant n)\times$$
$$\times\sum_{i=0}^{\infty}\mathbf{P}(\omega : \Delta_1(\omega)\neq n)^i=\mathbf{P}(\omega : S_k(\omega)=j_k,\ 0\leqslant k\leqslant n)/\mathbf{P}(\omega : \Delta_1(\omega)=n).$$

Отметим, что сл. в. $\tau_n(\omega)$ не является моментом остановки.]

11.14. Задача. Положим $\Lambda_{n,r}=\{\omega : S_n(\omega)=r\}$ при n и r одинаковой четности, $|r|\leqslant n$. Проверить, что условное блуждание $Y_k=(S_k\,|\,\Lambda_{n,r})$, $k\leqslant n$, — марковская цепь, выписать переходные вероятности и совместное распределение в точках $0<n_1<\ldots<n_k\leqslant\leqslant n$.

$$[\mathbf{P}(\omega : Y_k(\omega)=i_k,\ k\leqslant m)=\mathbf{P}(\omega : S_k(\omega)=i_k,\ k\leqslant m,\ S_n(\omega)=r)\times$$
$$\times\mathbf{P}(\omega : S_n(\omega)=r)^{-1}=N_{n-m,i_m-r}\cdot N_{n,r}^{-1},\quad N_{n,k}=C_n^{(n+k)/2},$$

где $m\leqslant n$, $|i_{k+1}-i_k|=1$, $k\leqslant m$, $i_0=0$, $i_n=r$. Отсюда получаем

$$p_{Y_{m+1}|Y_1,\ldots,Y_m}(i_{m+1}|i_1,\ \ldots,\ i_m)=N_{n-m-1,i_{m+1}-r}/N_{n-m,i_m-r}.$$

Переходные вероятности марковской цепи Y_k, $k\leqslant n$, равны

$$p_{Y_{m+1}\,|\,Y_m}(i\pm 1\,|\,i)=2^{-1}(1\mp(i-r)/(n-m))$$

при $\min(m,\ n-m+r)\geqslant i\geqslant\max(-m,\ -n+m+r)$. Аналогично

$$\mathbf{P}(\omega : Y_{n_i}(\omega)=m_i,\ i\leqslant k)=\prod_{i=1}^{k+1}N_{n_i-n_{i-1},m_i-m_{i-1}}\cdot N_{n,r}^{-1},$$

где $n_0=0<n_1<\ldots<n_k<n=n_{k+1}$, $m_0=0$, $m_{k+1}=r$.]

11.15. Задача. В условии 11.14 найти предельное распределение для конечномерных распределений процесса

$$(S_{[nt]}\,|\,\Lambda_{n,r})/\sqrt{n},\quad n\to\infty,\quad r/\sqrt{n}\to a.$$

[Из 11.14 выводим при $n_i=[nt_i]$, $m_i=[x_i\sqrt{n}]$, $i=1,\ \ldots,\ k$, $t_0==0<t_1<\ldots<t_k<1=t_{k+1}$ и любых $x_1,\ \ldots,\ x_k$, $(x_{k+1}=0)$:

$$\mathbf{P}(\omega : Y_{n_i}(\omega)=m_i,\ i\leqslant k)\sim(2/\sqrt{n})\,f_{t_1,\ldots,t_k,1}(x_1,\ \ldots,\ x_k,\ a)\,\varphi(a)^{-1},$$

где $f_{t_1,\ldots,t_l}$ есть плотность конечномерного распределения процесса броуновского движения $W(t)$ (ср. 11.10). Стандартный переход к интегральной теореме дает:

$$\mathbf{P}(\omega : Y_{n_i}(\omega)/\sqrt{n}\in I_{a_i,b_i},\quad i=1,\ \ldots,\ k)\to$$
$$\to\int_{a_1}^{b_1}\ldots\int_{a_k}^{b_k}f_{t_1,\ldots,t_k,1}(x_1,\ \ldots,\ x_k,\ a)\,\varphi(a)^{-1}\,dx_1,\ \ldots,\ dx_k.$$

Остается заметить, что подынтегральная функция совпадает с условной плотностью $f_{W(t_1),\ldots,W(t_k)\,|\,W(1)}(x_1, \ldots, x_k|a)$.]

11.16. Задача. Показать, что предельная конечномерная плотность в 11.15 представима в виде

$$f_{t_1,\ldots,t_k,1}(x_1, \ldots x_k, a)\varphi(a)^{-1}=\prod_{i=0}^{k-1} p_a(t_i, x_i, t_{i+1}, x_{i+1}), \quad 0<t_1<\ldots$$
$$\ldots<t_k<1, \quad t_0=x_0=0,$$

$$p_a(s, x, t, y)=\left(\frac{1-s}{(t-s)(1-t)}\right)^{1/2}\frac{\varphi((y-at)/\sqrt{1-t})}{\varphi((x-as)/\sqrt{1-s})}\times$$
$$\times\varphi\left(\frac{y-at-(x-as)}{\sqrt{t-s}}\right).$$

[Используя формулу конечномерной плотности броуновского движения (см. 11.10), запишем

$$f_{t_1,\ldots,t_k,1}(x_1, \ldots, x_k, a)/f_{t_1,\ldots,t_{k-1},1}(x_1, \ldots, x_{k-1}, a)=$$
$$=\left(\frac{1-t_{k-1}}{(t_k-t_{k-1})(1-t_k)}\right)^{1/2}\frac{\varphi((a-x_k)/\sqrt{1-t_k})}{\varphi((a-x_{k-1})/\sqrt{1-t_{k-1}})}\varphi\left(\frac{x_k-x_{k-1}}{t_k-t_{k-1}}\right).$$

Разделив полученное выражение на $p_a(t_{k-1}, x_{k-1}, t_k, x_k)$, покажем, что частное равно 1. Положим для краткости $t_{k-1}=s$, $t_k=t$, $x_{k-1}=x$, $x_k=y$. Логарифмируя указанное отношение и отбрасывая множитель 1/2, запишем

$$((a-y)^2-(at-y)^2)/(1-t)-((a-x)^2-(as-x)^2)/(1-s)+$$
$$+((y-x)^2-(y-at-(x-as))^2)/(t-s)=a^2(1+t)-2ay-$$
$$-(a^2(1+s)-2ax)+(2ay-2ax-a^2(t-s))=0,$$

что и требовалось установить.

Отметим, что поскольку

$$p_a(s, x, t, y)=p_0(s, x-as, t, y-at),$$

где $p_0(s, x, t, y)$ — переходная плотность броуновского моста (см. 11.10), то $p_a(s, x, t, y)$ является переходной плотностью. Марковский процесс $W_a(t)$, $0\leqslant t\leqslant 1$, с переходной плотностью $p_a(s, x, t, y)$ совпадает по распределению с $W_0(t)+at$ (ср. 11.11).]

11.17. Задача. Пользуясь формулой Стирлинга (см. § 3), показать, что равномерно по k, таким, что k, $n-k\to\infty$, имеет место соотношение

$$C_k^n\sim(2\pi n(k/n)(1-k/n))^{-1/2}\exp(nH(k/n)),$$
$$H(x)=-x\ln x-(1-x)\ln(1-x), \quad 0<x<1.$$

$$\Big[\, C_n^k = \frac{n!}{k!\,(n-k)!} \sim \frac{\sqrt{2\pi n}\, e^{-n} n^n}{\sqrt{2\pi k}\, e^{-k} k^k \sqrt{2\pi (n-k)}\, e^{-(n-k)} (n-k)^{n-k}} =$$

$$= \frac{\sqrt{n}}{\sqrt{2\pi k(n-k)}} \left(\left(\frac{k}{n}\right)^k \left(\frac{n-k}{n}\right)^{n-k} \right)^{-1} . \Big]$$

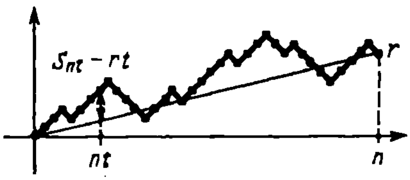

Рис. 31

11.18.* Задача. В условии 11.14 пусть $r=[n^\alpha]$ либо $r=[n^\alpha]+1$, $0<\alpha<1$, так, чтобы r было одной четности с n. Показать, что при $n\to\infty$ условный процесс

$$((S_{[nt]} - tr)/\sqrt{n} \mid \Lambda_{n,r}),$$
$$0 \leqslant t \leqslant 1,$$

сходится по распределению к броуновскому мосту (рис. 31).

[Для $Y_k = (S_k \mid \Lambda_{n,r})$ из 11.14 запишем условную вероятность (полагая $n(i) \equiv n_i$, $i=1, 2$):

$$p_{Y_{n(2)} \mid Y_{n(1)}} (m_2 \mid m_1) = N_{n_2-n_1, m_2-m_1} \cdot N_{n-n_2, r-m_2} \cdot N^{-1}_{n-n_1, r-m_1}.$$

Положим $n_i = [nt_i]$, $m_i = [rt_i + x_i\sqrt{n}\,]$, $i=1, 2$, $0<t_1<t_2<1$, и, воспользовавшись 11.17, запишем асимптотически эквивалентное при $n\to\infty$ выражение для этой вероятности:

$$C_{n_2-n_1}^{(n_2-n_1+m_2-m_1)/2} C_{n-n_2}^{(n-n_2+r-m_2)/2} \left(C_{n-n_1}^{(n-n_1+r-m_1)/2}\right)^{-1} \sim$$

$$\sim (2^{-1}\pi n(t_2-t_1))^{-1/2} \exp\left(n(t_2-t_1) H\left(2^{-1}\left(1 + n^{\alpha-1} + \frac{x_2-x_1}{t_2-t_1} n^{-1/2}\right)\right)\right) \times$$

$$\times (2^{-1}\pi n(1-t_2))^{-1/2} \exp\left(n(1-t_2) H\left(2^{-1}\left(1 + n^{\alpha-1} - \frac{x_2}{1-t_2} n^{-1/2}\right)\right)\right) \times$$

$$\times (2^{-1}\pi n(1-t_1))^{1/2} \exp\left(-n(1-t_1) H\left(2^{-1}\left(1 + n^{\alpha-1} - \frac{x_1}{1-t_1} n^{-1/2}\right)\right)\right).$$

Складывая показатели экспонент и используя разложение

$$H(2^{-1}(1+n^{\alpha-1}+\Delta n^{-1/2})) = H(2^{-1}(1+n^{\alpha-1})) + H'(2^{-1}(1+n^{\alpha-1})) \times$$
$$\times 2^{-1}\Delta n^{-1/2} + 2^{-1}H''(2^{-1}(1+n^{\alpha-1}))2^{-2}\Delta^2 n^{-1} + H'''(2^{-1}(1+n^{\alpha-1}+\theta\Delta n^{-1/2}))$$

при $\Delta=(x_2-x_1)/(t_2-t_1)$, $\Delta=-x_2/(1-t_2)$, $\Delta=-x_1/(1-t_1)$ и $0<\theta<1$, получим для их суммы выражение

$$n(t_2-t_1+1-t_2-1+t_1) H(2^{-1}(1+n^{\alpha-1})) + \sqrt{n}(x_2-x_1-x_2+x_1) \times$$
$$\times 2^{-1} H'(2^{-1}(1+n^{\alpha-1})) + 2^{-3}\{(x_2-x_1)^2/(t_2-t_1) + x_2^2/(1-t_2) -$$
$$- x_1^2/(1-t_1)\} H''(2^{-1}(1+n^{\alpha-1})) + o(1).$$

Поскольку коэффициенты при H и H' обращаются в нуль, а

$$H''(2^{-1}(1+n^{\alpha-1}))\to H''(2^{-1})=-4,\quad n\to\infty,$$

то получаем для условной вероятности эквивалентное выражение

$$\frac{2}{\sqrt{n}}\left(\frac{1-t_1}{(t_2-t_1)(1-t_2)}\right)^{1/2}\frac{\varphi(x_2/\sqrt{1-t_2})}{\varphi(x_1/\sqrt{1-t_1})}\varphi\left(\frac{x_2-x_1}{\sqrt{t_2-t_1}}\right),$$

что дает для вероятностей

$$\mathbf{P}(\omega: Y_{n_i}(\omega)=m_i,\ i\leqslant k),\quad n_i=[nt_i],\quad m_i=[rt_i+x_i\sqrt{n}]$$

такую же асимптотику, что и для соответствующих вероятностей в 11.10, откуда требуемый результат получается стандартным переходом к интегральной теореме.

Следует обратить внимание на то, что сл. в. $(S_{[nt]}|\Lambda_{n,r})$ в предельной теореме не только нормирована (множителем $1/\sqrt{n}$), но и *центрирована* вычитанием $tr=t[n^\alpha]$. При $\alpha<1/2$ $tr/\sqrt{n}\to 0$ и центрирование здесь использовано лишь для единообразия, при $\alpha=1/2$ оно сводится к постоянному сдвигу, однако при $\alpha>1/2$ оно существенно. В последнем случае распределение сл. в. $(S_{nt}|\Lambda_{n,r})$ не только «расползается» пропорционально $\sqrt{n}$, но и «сдвигается» на tn^α.]

11.19. Пример (гауссовские процессы). Конечномерные плотности броуновского движения и броуновского моста

$$f_{t_1,\ldots,t_k}(x_1,\ \ldots,\ x_k)=\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi(t_i-t_{i-1})}}\exp\left(-\frac{1}{2}\,\frac{(x_i-x_{i-1})^2}{t_i-t_{i-1}}\right),$$

$$f^0_{t_1,\ldots,t_k}(x_1,\ \ldots,\ x_k)=\sqrt{2\pi}f_{t_1,\ldots,t_k,1}(x_1,\ \ldots,\ x_k,\ 0)$$

являются многомерными нормальными плотностями (см. 9.18). Случайные процессы такого рода называются *нормальными* или *гауссовскими* в честь знаменитого немецкого математика К. Гаусса.

Нормальная плотность однозначно определяется матрицей ковариаций и вектором среднего. Чтобы задать гауссовский процесс $X(t)$, $t\in T\subseteq R$, надо задать его среднее $a(t)=\mathbf{M}X(t)$ и *ковариационную функцию* $r_0(s,\ t)=\mathbf{M}(X(s)-a(s))(X(t)-a(t))$, либо *корреляционную функцию* $r(s,\ t)=\mathbf{M}X(s)X(t)$. Броуновское движение и броуновский мост имеют нулевое среднее, гауссовский процесс $W^\alpha(t)$ из 11.16 имеет среднее at. Используя независимость приращений процесса броуновского движения $W(t)$ (см. 4.27 и 11.20), находим его корреляционную функцию

$$\mathbf{M}W(s)W(t)=\mathbf{M}W(s)(W(t)-W(s)+W(s))=$$
$$=\mathbf{M}W(s)(W(t)-W(s))+\mathbf{M}W(s)^2=0+s=s$$

при $s<t$, а при любых s, $t \geqslant 0$ имеем $r(s, t) = \min(s, t)$. Корреляционная функция броуновского моста находится из представления 11.11:

$$\mathbf{M}W^0(s)W^0(t) = \mathbf{M}(W(s) - sW(1))(W(t) - tW(1)) = \min(s, t) - st.$$

Конечно, эти результаты нетрудно получить непосредственно интегрированием

$$\iint xyf_{s,t}(x, y)\,dxdy.$$

11.20. Задача. Выразить совместную плотность приращений $W(t_1) = W(t_1) - W(0)$, $W(t_2) - W(t_1)$, ..., $W(t_k) - W(t_{k-1})$, $0 < t_1 < \ldots < t_k$, процесса броуновского движения (ср. 4.27).

[Положим $Y_i = W(t_i) - W(t_{i-1})$, $i = 1, \ldots, k$, $t_0 = 0$ и по формуле 3.26 запишем

$$f_{Y_1,\ldots,Y_k}(y_1, \ldots, y_k) = f_{W(t_1),\ldots,W(t_k)}(x_1, \ldots, x_k),$$

где переменные $y_1, \ldots, y_k$ и $x_1, \ldots, x_k$ связаны линейными соотношениями с единичным детерминантом

$$y_i = x_i - x_{i-1}, \quad i = 1, \ldots, k, \; x_0 = 0.$$

Отсюда находим (см. 11.19)

$$f_{Y_1,\ldots,Y_k}(y_1, \ldots, y_k) = \prod_{i=1}^{k} \left(1/\sqrt{t_i - t_{i-1}}\right) \varphi\left(y_i/\sqrt{t_i - t_{i-1}}\right).$$

т. е. процесс $W(t)$ имеет независимые и *стационарные* или *однородные по времени* приращения.]

11.21.* Задача. В условии 11.18 заменить n^α на θn, а $0<\theta<1$, и показать, что предельный процесс есть $W^0(t) \cdot \sqrt{1-\theta^2}$.

[Действуя аналогично 11.18, получаем для $Y_k = (S_k \mid \Lambda_{n,r})$, $n_i \equiv n(i) = [nt_i]$, $m_i = \left[rt_i + x_i\sqrt{n}\right]$, $i = 1, 2$, $0 < t_1 < t_2 < 1$:

$$p_{Y_{n(2)} \mid Y_{n(1)}}(m_2 \mid m_1) \sim \frac{2}{\sqrt{n}} \left(\frac{1 - t_1}{(t_2 - t_1)(1 - t_2)}\right)^{1/2} \cdot \frac{(1-\theta^2)^{-1/2}}{\sqrt{2\pi}} \times$$

$$\times \exp\{n(t_2 - t_1) H((1+\theta)2^{-1} + 2^{-1}n^{-1/2}(x_2 - x_1)/(t_2 - t_1)) \times$$

$$\times n(1 - t_2) H((1+\theta)2^{-1} - 2^{-1}n^{-1/2}x_2/(1 - t_2)) \times$$

$$\times (-n(1 - t_1) H((1+\theta)2^{-1} - 2^{-1}n^{-1/2}x_1/(1 - t_1))\}.$$

Используя разложение по формуле Тейлора

$$H((1+\theta)2^{-1} + 2^{-1}n^{-1/2}\Delta) = H((1+\theta)2^{-1}) - H'((1+\theta)2^{-1}) \times$$

$$\times 2^{-1}n^{-1/2}\Delta + 2^{-1}H''((1+\theta)2^{-1})2^{-2}n^{-1}\Delta^2 + o(n^{-1}),$$

перепишем выражение под знаком экспоненты в виде

$$H((1+\theta)2^{-1})((t_2-t_1)+(1-t_2)-(1-t_1))n+$$
$$+H'((1+\theta)2^{-1})(2^{-1}(x_2-x_1)-2^{-1}x_2-2^{-1}x_1)\sqrt{n}+o(1)+$$
$$+2^{-1}H''((1+\theta)2^{-1})2^{-2}((x_2-x_1)^2/(t_2-t_1)+x_2^2/(1-t_2)-x_1^2/(1-t_1)).$$

Учитывая, что множители при $H((1+\theta)2^{-1})$ и $H'((1+\theta)2^{-1})$ равны нулю, а $H''((1+\theta)2^{-1})=-2^2(1-\theta^2)^{-1}$, получаем для условной вероятности асимптотически эквивалентное выражение

$$\frac{2}{\sqrt{n}}\left(\frac{1-t_1}{(t_2-t_1)(1-t_2)}\right)^{1/2}\frac{(1-\theta^2)^{-1/2}}{\sqrt{2\pi}}\exp\left(-\frac{(1-\theta^2)^{-1}}{2}\frac{(x_2-x_1)^2}{t_2-t_1}\right)\times$$
$$\times\exp\left(-\frac{(1-\theta^2)^{-1}}{2}\left(\frac{x_2^2}{1-t_2}-\frac{x_1^2}{1-t_1}\right)\right)=$$
$$=\frac{2}{\sqrt{n}}\frac{\sqrt{(1-t_1)(1-\theta^2)}}{\sqrt{(t_2-t_1)(1-\theta_2)}\sqrt{(1-t_2)(1-\theta^2)}}\varphi\left(\frac{x_2-x_1}{\sqrt{(t_2-t_1)(1-\theta^2)}}\right)\times$$
$$\times\varphi\left(\frac{x_2}{\sqrt{(1-t_2)(1-\theta^2)}}\right)\Big/\varphi\left(\frac{x_1}{\sqrt{(1-t_1)(1-\theta^2)}}\right).$$

Переход к интегральной теореме для конечномерных распределений показывает, что предельный процесс есть $W^0(t)\sqrt{1-\theta^2}$.]

11.22.* Задача. Положим $\Lambda_n^+=\{\omega: S_k(\omega)>0,\ 0<k<n,\ S_n(\omega)=0\}$ при четном n. Показать, что условное случайное блуждание $Y_k=(S_k|\Lambda_n^+)$, $0\leqslant k\leqslant n$, — марковская цепь, выписать ее переходные вероятности и совместное распределение вероятностей в точках $0<n_1<\ldots<n_k<n$. Получить предельную теорему по образцу 11.10.

[Обозначим $N_{n,k}^+$, $k\geqslant 0$, число путей, ведущих из точки (0, 0) в точку (n, k) и лежащих целиком выше оси абсцисс, не считая начальной и конечной точек. Поскольку имеется столько же путей длины n, достигающих в момент n точку k впервые (см. 1.16), то (см. 11.7)

$$N_{n,k}^+=(k/n)C_n^{(n+k)/2}=(k/n)N_{n,k},\quad k>0,\quad N_{n,0}^+\equiv N_{n-1,1}^+$$

при $k\leqslant n$, k одной четности с n. Отсюда при $m\leqslant n$ $|i_{k+1}-i_k|=1$, $k\leqslant m$, $i_k>0$, $0<k<n$, $i_0=i_n=0$, получаем

$$\mathbf{P}(\omega: Y_k(\omega)=i_k,\ k\leqslant m)=N_{n-m,i_m}^+/N_{n,0}^+,$$

$$p_{Y_{m+1}|Y_1,\ldots,Y_m}(i_{m+1}|i_1,\ \ldots,\ i_m)=N_{n-m-1,i_{m+1}}^+/N_{n-m,i_m}^+,$$

$$p_{Y_{m+1}|Y_m}(i\pm 1|i)=2^{-1}(1\pm i^{-1})(1\mp i/(n-m)),\quad 0<i\leqslant\min(m,\ n-m).$$

Число положительных путей, ведущих из точки $(0, a)$, $a>0$, в точку (n, b), $b>0$, равно (см. 1.14) $N_{n,\,b-a}-N_{n,\,b+a}$, откуда при $0<n_1<\ldots<n_k<n$, $m_i>0$, $i=1,\ldots,k$, имеем для совместного распределения выражение

$$\mathbf{P}(\omega : Y_{n_i}(\omega)=m_i,\ i\leqslant k)=(N^{+}_{n,0})^{-1} N^{+}_{n_1,m_1} N^{+}_{n-n_k,m_k}\times$$
$$\times\prod_{i=2}^{k}(N_{n_i-n_{i-1},m_i-m_{i-1}}-N_{n_i-n_{i-1},m_i+m_{i-1}}).$$

Подставляя $n_i=[nt_i]$, $m_i=[x_i\sqrt{n}]$, $0<t_1<\ldots<t_k<1$, $x_i>0$, $i=1,\ldots,k$, получаем, как и в 11.10, асимптотически эквивалентное выражение для совместного распределения

$$n^{3/2}\sqrt{\frac{\pi}{2}}\,\frac{x_1}{t_1}\,n^{-1/2}\,\frac{2}{\sqrt{nt_1}}\,\varphi\left(\frac{x_1}{\sqrt{t_1}}\right)\cdot\frac{x_k}{1-t_k}\,n^{-1/2}\,\frac{2}{\sqrt{n(1-t_k)}}\times$$
$$\times\varphi\left(\frac{x_k}{\sqrt{1-t_k}}\right)\prod_{i=2}^{k}\frac{2}{\sqrt{n}}\,\frac{1}{\sqrt{t_i-t_{i-1}}}\times$$
$$\times\left(\varphi\left(\frac{x_i-x_{i-1}}{\sqrt{t_i-t_{i-1}}}\right)-\varphi\left(\frac{x_i+x_{i-1}}{\sqrt{t_i-t_{i-1}}}\right)\right).$$

Стандартные рассуждения приводят к интегральной теореме. Конечномерные плотности предельного процесса равны

$$f_{t_1,\ldots,t_k}(x_1,\ \ldots\ ,\ x_k)=\sqrt{2\pi}\,x_1t_1^{-3/2}\varphi\left(\frac{x_1}{\sqrt{t_1}}\right)x_k(1-t_k)^{-3/2}\varphi\left(\frac{x_k}{\sqrt{1-t_k}}\right)\times$$
$$\times\prod_{i=2}^{k}\left(1/\sqrt{t_i-t_{i-1}}\right)\left(\varphi\left((x_i-x_{i-1})/\sqrt{t_i-t_{i-1}}\right)-\right.$$
$$\left.-\varphi\left((x_i+x_{i-1})/\sqrt{t_i-t_{i-1}}\right)\right).$$

Процесс с конечномерными плотностями $f_{t_1,\ldots,t_k}$ называется *броуновской экскурсией* и обозначается $W_0^{+}(t)$, $0\leqslant t\leqslant 1$. Как видно, $W_0^{+}(t)$ — марковский, его переходная плотность равна при $0<s<t<1$, $x>0$, $y>0$:

$$p_0^{+}(s,\ x,\ t,\ y)=\frac{y(1-t)^{-3/2}\varphi(y/\sqrt{1-t})}{x(1-s)^{-3/2}\varphi(x/\sqrt{1-s})}\cdot\frac{1}{\sqrt{t-s}}\times$$
$$\times\left(\varphi\left(\frac{y-x}{\sqrt{t-s}}\right)-\varphi\left(\frac{y+x}{\sqrt{t-s}}\right)\right).\ ]$$

§ 12.

СХЕМА БЕРНУЛЛИ И ПРОСТОЕ БЛУЖДАНИЕ

12.1. Пример (биномиальное распределение). Исследование биномиального распределения

$$p_n(k;\ p)=C_n^k p^k(1-p)^{n-k},\quad k=0,\ 1,\ \ldots,\ n,$$

можно начать с поведения $p_n(k;\ p)$ как функции от k. Отношение

$$\frac{p_n(k;\ p)}{p_n(k-1;\ p)}=\frac{n!\,p^k(1-p)^{n-k}(k-1)!\,(n-k+1)!}{k!\,(n-k)!\,n!\,p^{k-1}(1-p)^{n-k+1}}=$$
$$=\frac{(n-k+1)}{k(1-p)}=1+\frac{(n+1)p-k}{k(1-p)}$$

больше единицы при $k<(n+1)p$ и меньше единицы при $k>$ $>(n+1)p$. Таким образом, вероятности $p_n(k;\ p)$ возрастают при изменении k от 0 до $[(n+1)p]$ и убывают для k от $[(n+1)p]+1$ до n. Максимальное значение достигается в точке $k=[(n+1)p]$, если число $(n+1)p$ нецелое, в противном случае наибольшее значение принимается дважды: в точках $(n+1)p-1$ и $(n+1)p$. Характер изменения вероятностей $p_n(k,\ p)$ при $n\to\infty$ удобно изучать с помощью формулы (см. 11.17)

$$C_n^k\sim(2\pi n(k/n)(1-k/n))^{-1/2}\exp(nH(k/n)),$$
$$H(x)=-x\ln x-(1-x)\ln(1-x),\quad 0<x<1,$$

при $k,\ n-k\to\infty$. Обозначив

$$H(x;\ p)=-x\ln p-(1-x)\ln(1-p)=$$
$$=x(-\ln p+\ln(1-p))-\ln(1-p),\quad 0<p<1,$$

запишем асимптотическое выражение для биномиальных вероятностей при $k,\ n-k\to\infty$ в виде

$$p_n(k;\ p)\sim\left(2\pi n\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{-1/2}\exp\left(-n\left(H\left(\frac{k}{n};\ p\right)-H\left(\frac{k}{n}\right)\right)\right).$$

Так как

$$H'(x)=-\ln x+\ln(1-x),\ H(p;\ p)=H(p),\ H'(x;\ p)|_{x=p}=H'(p),$$

то $H(x;\ p)$ представляет собой касательную к графику $H(x)$ в точке $x=p$ (рис. 32). При k, для которых отношение k/n отделено от p, порядок убывания вероятностей $p_n(k;\ p)$ с ростом n — экспоненциальный с показателем в экспоненте тем большим, чем дальше расположено k/n от p. Для $k/n \to p$ воспользуемся разложением функции $H(x)$ в ряд Тейлора

$$H(x) = H(p) + H'(p)(x-p) + H''(p)(x-p)^2/2! + \dots$$

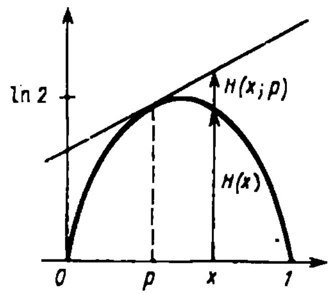

Рис. 32

и запишем

$$H\left(\frac{k}{n};\ p\right) - H\left(\frac{k}{n}\right) = -\frac{1}{2}H''(p)\left(\frac{k}{n}-p\right)^2 = -\frac{1}{3!}H'''(p)\left(\frac{k}{n}-p\right)^3 - \dots .$$

При $n(k/n-p)^3 \to 0$, т. е. $k-np = o(n^{2/3})$, вторым членом этого разложения можно пренебречь и, принимая во внимание, что $H''(x) = -(x(1-x))^{-1}$, записать соотношение

$$p_n(k;\ p) \sim (2\pi np(1-p))^{-1/2} \exp\left(-\frac{1}{2}\,\frac{(k-np)^2}{np(1-p)}\right)$$

равномерно по k, таким, что $k-np = o(n^{2/3})$. Это так называемая *локальная теорема Муавра—Лапласа*. При $k/n - p = o(n^{(m-1)/m})$ и любом натуральном $m \geqslant 3$ достаточно для получения асимптотики вероятностей $p_n(k;\ p)$ взять в написанном выше разложении члены до m-й производной включительно. Отметим, что для $p=1/2$ $H'''(1/2)=0$, и потому локальная теорема действует в области $k-np=o(n^{3/4})$ (ср. 3.18).

Суммируя вероятности $p_n(k;\ p)$ по $(k-np)/\sqrt{np(1-p)} \in [a,\ b]$, $-\infty < a < b < \infty$, и вводя последовательность вероятностных мер

$$\mathbf{P}_n^*(I_{a,b}) = \sum_{\{(k-np)/\sqrt{np(1-p)} \in I_{a,b}\}} p_n(k;\ p),$$

запишем *интегральную теорему Муавра—Лапласа*

$$\mathbf{P}_n^*(I_{a,b}) \to \int_a^b \varphi(x)\,dx,\quad \varphi(x) = \left(1/\sqrt{2\pi}\right)e^{-x^2/2}.$$

Простые рассуждения (ср. 3.15) позволяют распространить сходимость на любой промежуток $I_{a,b}$, $-\infty \leqslant a < b \leqslant \infty$.

В области $k-np = O\left(\sqrt{n}\right)$ биномиальные вероятности $p_n(k;p)$ имеют одинаковый порядок малости при $n \to \infty$. За пределами отрезка $\left[np - A_\varepsilon\sqrt{n},\ np + A_\varepsilon\sqrt{n}\right]$ расположена биномиальная масса $<\varepsilon$ при достаточно большом A_ε и всех n, начиная с некоторого. Область $k-np = O\left(\sqrt{n}\right)$ называют областью *нормальных уклонений* (значений биномиальной случайной величины от своего среднего). Значения k, такие, что $|k-np|/\sqrt{n}\to\infty$, называют *большими уклонениями.*

Поскольку асимптотика

$$p_n(k;\ p)\sim(np(1-p))^{-1/2}\,\varphi\left((k-p)/\sqrt{np(1-p)}\right)$$

действует равномерно в области $(k-np)/\sqrt{np(1-p)}=o(n^{1/6})$, то, как легко видеть, при $b_n-a_n\geqslant\varepsilon>0$

$$\mathbf{P}_n^*(I_{a_n,b_n})\sim\left(1/\sqrt{2\pi}\right)\int\limits_{a_n}^{b_n}e^{-x^2/2}\,dx$$

при любых a_n, $b_n=o(n^{1/6})$, $n\to\infty$. Если $a_n\to-\infty$, $b_n\to\infty$, то интеграл справа стремится к 1 и соотношение не прибавляет ничего нового. Однако если a_n, b_n стремятся к бесконечности одного знака, то это соотношение приводит к содержательным выводам. Ввиду очевидной симметрии задачи ограничим рассмотрение случаем положительных a_n, b_n. При $\lambda\to\infty$ и любом $\varepsilon>0$ (см. ниже 12.2)

$$\int\limits_{\lambda}^{\infty}e^{-x^2/2}\,dx\sim\int\limits_{\lambda}^{\lambda+\varepsilon}e^{-x^2/2}\,dx\sim\lambda^{-1}e^{-\lambda^2/2}.$$

Это означает, что при условии $b_n-a_n\geqslant\varepsilon>0$ асимптотика биномиальной меры отрезка $\left[np+a_n\sqrt{np(1-p)},\ np+b_n\sqrt{np(1-p)}\right]$ от b_n не зависит и

$$\mathbf{P}_n^*([a_n,\ b_n])\sim\int\limits_{a_n}^{\infty}\varphi(x)\,dx\sim a_n^{-1}\varphi(a_n).$$

Покажем, что при $a_n=o(n^{1/6})$, $a_n>0$, имеет место соотношение

$$\mathbf{P}_n^*([a_n,\ \infty))\sim\int\limits_{a_n}^{\infty}\varphi(x)dx,\quad n\to\infty.$$

Запишем

$$\mathbf{P}_n^*([a_n,\ \infty))=\mathbf{P}_n^*([a_n,\ a_n+n^{1/7}))+\mathbf{P}_n^*([a_n+n^{1/7},\ \infty)),$$

где первое слагаемое в правой части эквивалентно при $n\to\infty$ $a_n^{-1}\varphi(a_n)$, а второе оценивается сверху величиной

$$\mathbf{P}_n^*([a_n+n^{1/7},\ \infty))=\sum_{\left\{k\geqslant np+(a_n+n^{1/7})\sqrt{np(1-p)}\right\}}p_n(k,\ p)\leqslant$$

$$\leqslant(n-np)\,p_n\left(\left[np+(a_n+n^{1/7})\sqrt{np(1-p)}\right];\ p\right)\sim(n(1-p)/p)^{1/2}\times$$
$$\times(2\pi)^{-1/2}\exp\left(-2^{-1}(a_n^2+2a_nn^{1/7}+n^{2/7})\right)\leqslant(n(1-p)/p)^{1/2}\,\varphi(a_n)\,e^{-n^{2/7}/2}$$

более высокого порядка малости, чем $a_n^{-1}\varphi(a_n)$.

В области больших уклонений *хвосты распределений вероятностей* $\mathbf{P}_n^*([x_n, \infty))$, $\mathbf{P}_n^*((-\infty, -x_n])$, $x_n \to \infty$, представляют наибольший интерес. При $x_n = o(n^{1/6})$, как мы видим, они *следуют* за хвостами нормального распределения. В более далекой области при $p \neq 1/2$ асимптотика у хвостов распределения $\mathbf{P}_n^*$ уже иная. Для очень больших уклонений (x_n порядка $\sqrt{n}$, т. е. $k-np$ порядка n) асимптотическая формула получена в 12.4. В промежутке между $n^{1/6}$ и $n^{1/2}$ соответствующие формулы более громоздки.

Для иллюстрации оценочного значения приближений, даваемых локальной и интегральной теоремами Муавра—Лапласа, в табл. 1 приведены значения биномиальных вероятностей

Таблица 1

k	$p_n(k;\ 1/2)$	$(2/\sqrt{n})\ \varphi(x_k)$	$\mathbf{P}_n^*((-\infty,\ x_k])$	$\Phi(x_k + 1/\sqrt{n})$
8	0,0054	0,0055	0,0080	0,0088
9	0,0133	0,0132	0,0214	0,0223
10	0,0279	0,0275	0,0493	0,0501
11	0,0508	0,0502	0,1002	0,1005
12	0,0805	0,0800	0,1808	0,1807
13	0,1115	0,1115	0,2923	0,2920
14	0,1354	0,1362	0,4277	0,4276
15	0,1444	0,1456	0,5722	0,5724

$p_n(k;\ p)$, $\mathbf{P}_n^*((-\infty, x_k])$, $x_k = (k-np)/\sqrt{np(1-p)}$, и их нормальные приближения $\left(1/\sqrt{np(1-p)}\right)\varphi(x_k)$, $\Phi\left(x_k + 1/\left(2\sqrt{np(1-p)}\right)\right)$, где Φ есть функция распределения стандартного нормального закона (поправка на дискретность $1/\left(2\sqrt{np(1-p)}\right)$ улучшает точность приближения при умеренных значениях n), $p = 1/2$, $n = 30$.

Для использования теоремы Муавра—Лапласа для конечных n необходимо заметить, что погрешность на самом деле определяется значением $np(1-p)$ дисперсии распределения. Остаточный член в экспоненте локальной теоремы Муавра—Лапласа

$$-\frac{1}{3!}\,nH'''(p)\,(k/n-p)^3 = \frac{1}{3!}\,(1-2p)\,(k-np)^3/(np\,(1-p))^2$$

зависит от n через значение $np(1-p)$ и отклонение $k-np$, откуда следует, что формула нормального приближения для биномиальных вероятностей $p_n(k;\ p)$ остается справедливой и при $p = p_n \to 0$ или $p = p_n \to 1$, если только $np(1-p) \to \infty$, а $k - np = o\left(np(1-p)\right)^{2/3}$. Приближение $\mathbf{P}_n^*((-\infty, x_k]) \approx \Phi\left(x_k + 1/\left(2\sqrt{np(1-p)}\right)\right)$ оказывается неплохим даже при весьма умеренных значениях $np(1-p)$. В табл. 2 взяты $n = 30$, $p = 0{,}2$, так что $np = 6$, $np(1-p) = 4{,}8$.

Таблица 2

k	2	4	6	8	10
$\mathbf{P}_n^*((-\infty, x_k])$	0,0441	0,2552	0,6069	0,8713	0,9748
$\Phi(x_k+1/(0{,}8\sqrt{n}))$	0,0551	0,2468	0,5902	0,8730	0,9800

Наконец, если $p=p_n\to 0$ при $n\to\infty$, но $np_n(1-p_n)\to\lambda>0$, то биномиальное распределение приближается пуассоновским (см. 3.11):

$$p_n(k;\ p)\to(\lambda^k/k!)e^{-\lambda}=p(k;\ \lambda),\quad k=0,\ 1,\ 2,\ \ldots$$

(и аналогичная формула для вероятностей $p_n(n-k;\ p)$ при $p=$ $=p_n\to 1$, $np_n(1-p_n)\to\lambda$). В табл. 3 сравниваются вероятности $p_n(k;\ p)$ и $p(k;\ \lambda)$ для $n=30$, $p=0{,}1$, $\lambda=np=3$.

Таблица 3

k	0	2	4	6	8
$p_n(k;\ p)$	0,0424	0,2276	0,1770	0,0473	0,0057
$p_n(k;\ \lambda)$	0,0498	0,2240	0,1680	0,0504	0,0081

Биномиальное распределение — распределение вероятностей сл. в. $S_n=X_1+\ldots+X_n$, где X_i, $i=1, \ldots, n$, — независимые случайные величины, принимающие значения 1 и 0 с вероятностями p и $1-p$ соответственно. Интегральную теорему Муавра—Лапласа можно записать в виде

$$\mathbf{P}\left(\omega : (S_n(\omega)-\mathbf{M}S_n)/\sqrt{\mathbf{D}S_n}\leqslant x\right)\to\Phi(x),\ n\to\infty.$$

Замечательное положение теории вероятностей состоит в том, что это соотношение сохраняется и тогда, когда $X_1, \ldots, X_n$ — произвольные независимые случайные величины с одинаковым распределением и конечной дисперсией (а также и в значительно более общей ситуации). С некоторыми частными случаями этой центральной предельной теоремы встретимся ниже, а в общем виде она будет установлена в § 13.

12.2. Задача. Показать, что

$$\mathcal{J}(\lambda)\equiv\int\limits_{\lambda}^{\infty}e^{-x^2/2}dx\sim\lambda^{-1}e^{-\lambda^2/2},\quad \lambda\to\infty,$$

установив с помощью замены переменного $x=\lambda+t$, что:

$$\text{(I)}\ \mathcal{J}(\lambda)\leqslant\lambda^{-1}e^{-\lambda^2/2};\quad \text{(II)}\ \mathcal{J}(\lambda)-\mathcal{J}(\lambda+\varepsilon)\geqslant e^{-\varepsilon^2/2-\lambda^2/2}\int_0^\varepsilon e^{-\lambda t}\,dt.$$

$$\Big[\ \mathcal{J}(\lambda)=\int_\lambda^\infty e^{-x^2/2}dx=\int_0^\infty e^{-\lambda^2/2-\lambda t-t^2/2}\,dt\leqslant e^{-\lambda^2/2}\int_0^\infty e^{-\lambda t}\,dt=\lambda^{-1}e^{-\lambda^2/2}.$$

$$\mathcal{J}(\lambda)-\mathcal{J}(\lambda+\varepsilon)=\int_\lambda^{\lambda+\varepsilon}e^{-x^2/2}\,dx=\int_0^\varepsilon e^{-\lambda^2/2-\lambda t-t^2/2}\,dt\geqslant$$

$$\geqslant e^{-\varepsilon^2/2-\lambda^2/2}\int_0^\varepsilon e^{-\lambda t}\,dt=e^{-\varepsilon^2/2-\lambda^2/2}(1-e^{-\lambda\varepsilon})\lambda^{-1},\quad \varepsilon>0.$$

Замечая, что $\mathcal{J}(\lambda)\geqslant\mathcal{J}(\lambda)-\mathcal{J}(\lambda+\varepsilon)$, получаем из (II):

$$\varliminf_{\lambda\to\infty}\mathcal{J}(\lambda)/(\lambda^{-1}e^{-\lambda^2/2})\geqslant e^{-\varepsilon^2/2}.$$

Устремляя ε к 0, имеем

$$\varliminf_{\lambda\to\infty}\mathcal{J}(\lambda)/(\lambda^{-1}e^{-\lambda^2/2})\geqslant 1.$$

Принимая во внимание (I), получаем требуемый результат.]

12.3. Задача. Вывести оценки хвостов биномиального распределения:

$$\sum_{k=0}^m p_n(k;\ p)\leqslant p_n(m;\ p)\frac{(n-m+1)p}{(n+1)p-m},\quad m<(n+1)p,$$

$$\sum_{k=m}^n p_n(k;\ p)\leqslant p_n(m;\ p)\frac{(m+1)(1-p)}{(n+1)(1-p)-n+m},\quad m>(n+1)p.$$

[Имеем при $k<m$

$$C_n^k/C_n^m=\prod_{i=k}^{m-1}C_n^i/C_n^{i+1}=\frac{k+1}{n-k}\cdot\frac{k+2}{n-k+1}\cdot\ldots\cdot\frac{m}{n-m+1}\leqslant\left(\frac{m}{n-m+1}\right)^{m-k},$$

$$\sum_{k=0}^m p_n(k;\ p)\leqslant p_n(m;\ p)\sum_{k=0}^m p_n(k;\ p)/p_n(m;\ p)\leqslant$$

$$\leqslant p_n(m;\ p)\sum_{k=0}^m\left(\frac{m}{n-m+1}\cdot\frac{1-p}{p}\right)^{m-k}.$$

Заменяя конечную геометрическую прогрессию бесконечной и учитывая, что при $m<(n+1)p$

$$1-\frac{m}{n-m+1}\cdot\frac{1-p}{p}=\frac{(n-m+1)p}{(n+1)p-m}>0,$$

получаем первое из требуемых неравенств. Второе неравенство вытекает из первого после замены p на $1-p$ ввиду равенства

$$p_n(k;\ 1-p)=p_n(n-k;\ p).]$$

12.4. Задача. Вывести асимптотику хвоста распределения биномиальной случайной величины:

$$\mathbf{P}(\omega:S_n(\omega)\geqslant[n\theta])\sim p_n([n\theta];\ p)\,\theta(1-p)/(\theta-p),\quad p<\theta<1.$$

[Из неравенств 12.3 получаем

$$\varlimsup_{n\to\infty} p_n([n\theta];\ p)^{-1}\sum_{\{k\geqslant n\theta\}} p_n(k;\ p)\leqslant\theta(1-p)/(\theta-p).$$

С другой стороны, используя в рассуждениях 12.3 вместо оценки сверху аналогичную оценку снизу при $k>m$

$$C_n^k/C_n^m=\frac{n-k+1}{k}\cdot\frac{n-k+2}{k-1}\cdot\ldots\cdot\frac{n-m}{m+1}\geqslant\left(\frac{n-k+1}{k}\right)^{k-m},$$

запишем при $0<r<n(1-\theta)$

$$\sum_{\{k\geqslant n\theta\}} p_n(k;\ p)\geqslant p_n([n\theta];\ p)\sum_{k=[n\theta]}^{[n\theta]+r}\left(\frac{n-k+1}{k}\cdot\frac{p}{1-p}\right)^{k-[n\theta]}\geqslant$$

$$\geqslant p_n([n\theta];\ p)\sum_{k=[n\theta]}^{[n\theta]+r}\left(\frac{n-n\theta-r+1}{n\theta+r}\cdot\frac{p}{1-p}\right)^{k-[n\theta]},$$

где учтено, что отношение $(n-k+1)/k$ убывает по k. Переходя к пределу при фиксированном r, получаем

$$\varliminf_{n\to\infty} p_n([n\theta];\ p)^{-1}\sum_{\{k\geqslant n\theta\}} p_n(k;\ p)\geqslant\sum_{k=0}^{r}\left(\frac{1-\theta}{\theta}\cdot\frac{p}{1-p}\right)^{k}.$$

Устремляя $r\to\infty$, находим, что нижняя оценка совпадает с верхней.]

12.5. Задача. Показать, что последовательность пуассоновских распределений вероятностей с параметром $n=1, 2, \ldots$

$$p(k;\ n)=(n^k/k!)e^{-n},\quad k=0,\ 1,\ 2,\ \ldots,$$

после надлежащего центрирования и нормирования сходится к нормальному закону, т. е. для некоторых $\beta_n>0$, α_n и

$$\mathbf{P}_n^*(I_{a,b})=\sum_{\{(k-\alpha_n)/\beta_n\in I_{a,b}\}} p(k;\ n)$$

имеет место соотношение

$$\mathbf{P}_n^*(I_{a.b})\to\int_a^b \varphi(x)\,dx,\quad n\to\infty.$$

[Отношение пуассоновских вероятностей

$$p(k;\ n)/p\,(k-1;\ n)=n/k$$

достигает единицы в точке $k=n$, так что максимальное значение принимается дважды и равно

$$p(n;\ n)=(n^n/n!)\,e^{-n}\sim(2\pi n)^{-1/2},\quad n\to\infty,$$

где применена формула Стирлинга. Исследуем вероятности $p(k;\ n)$ в окрестности максимума $k-n=o(n)$. При $k>n$

$$\frac{p\,(k;\ n)}{p\,(n;\ n)}=\prod_{i=1}^{k-n}\frac{p\,(n+i;\ n)}{p\,(n+i-1;\ n)}=\prod_{i=1}^{k-n}\left(1+\frac{i}{n}\right)^{-1},$$

а при $k<n$

$$p(k;\ n)/p(n;\ n)=\prod_{i=0}^{n-k-1}(1-i/n).$$

Разлагая логарифм по формуле Тейлора до второго члена, получаем

$$\ln\prod_{i=1}^{|k-n|}(1\pm i/n)=\sum_{i=1}^{|k-n|}\ln(1\pm i/n)=\sum_{i=1}^{|k-n|}(\pm i/n+O(1)\,(i/n)^2),$$

где множитель $O(1)$ равномерно ограничен в области $k-n=o(n)$. Принимая во внимание, что

$$1+2+\ldots+m=m(m+1)/2,\quad 1^2+2^2+\ldots+m^2=O(m^3),$$

находим

$$p(k;\ n)\sim(2\pi n)^{-1/2}\exp\left(-(k-n)^2/(2n)+O\left((k-n)^3/n^2\right)\right).$$

В области $(k-n)^3=o(n^2)$, т. е. $k-n=o(n^{2/3})$, получаем равномерно по k

$$p(k;\ n)\sim(2\pi n)^{-1/2}\exp\left(-(k-n)^2/(2n)\right).$$

Стандартный переход к интегральной теореме для $\mathbf{P}_n^*$ с $\alpha_n = n$, $\beta_n = \sqrt{n}$ заканчивает доказательство.

Отметим, что сл. в. $S_n = X_1 + \ldots + X_n$, где X_i, $i = 1, 2, \ldots$, — независимые одинаково распределенные по Пуассону с параметром единица, имеет пуассоновское распределение с параметром n (см. 4.7), $\mathbf{M}X_1 = \mathbf{D}X_1 = 1$ (см. 7.2, 8.7), так что $\mathbf{P}_n^*$ — распределение вероятностей *стандартизованной* сл. в. $(S_n - \mathbf{M}S_n)/\sqrt{\mathbf{D}S_n}$. Полученный результат — частный случай центральной предельной теоремы, о которой упоминалось в 12.1.

При аппроксимации биномиального распределения пуассоновским (при небольших $np(1-p)$) и нормальным (при больших) в связи с установленным результатом заключаем, что для промежуточных значений $np(1-p)$ мы не получим противоречия, если будем пользоваться любым из этих двух приближений.]

12.6. Задача. Обозначим τ_r число неудач, предшествующих r-му успеху в последовательности испытаний Бернулли. Установить, что после надлежащего центрирования и нормирования распределения вероятностей сл. в. τ_r сходятся к нормальному при $r \to \infty$: (I) пользуясь тем, что τ_r имеет отрицательное биномиальное распределение (см. 2.17)

$$p_{\tau_r}(k) = C_{k+r-1}^k (1-p)^k p^r, \quad k = 0, 1, 2, \ldots,$$

и что $p_{\tau_r}(k) = p \cdot p_{r+k-1}(k;\ 1-p)$, где $p_n(k;\ p)$ обозначает биномиальное распределение с параметрами n, p; (II) используя соотношение $\tau_r > k \Leftrightarrow S_{k+r} < r$, где S_n — число успехов в n испытаниях Бернулли.

[(I) Из 12.1 получаем

$$p_{\tau_r}(k) \sim p\,(2\pi(r+k)\,p\,(1-p))^{-1/2} \exp\left(-\frac{1}{2}\,\frac{(k-(r+k)(1-p))^2}{(r+k)\,p\,(1-p)}\right)$$

при $(r+k)\,p\,(1-p) \to \infty$, $k - (r+k)(1-p) = o\,((r+k)\,p\,(1-p))^{2/3}$. Преобразовав экспоненту в правой части к виду

$$\exp\left(-\frac{1}{2}(k - r(1-p)/p)^2\,((r+k)(1-p)/p)^{-1}\right),$$

рассмотрим k, такие, что $k - r(1-p)/p = O(\sqrt{r})$, $r \to \infty$, предполагая, что вероятность успеха $0 < p < 1$ фиксирована. В таком случае

$$k \sim r(1-p)/p, \quad k + r \sim r/p, \quad (r+k)\,p\,(1-p) \sim r(1-p),$$
$$(r+k)(1-p)/p \sim r(1-p)/p^2, \quad r \to \infty,$$

откуда получаем равномерно по k

$$p_{\tau_r}(k) \sim (2\pi r\,(1-p)/p^2)^{-1/2} \exp\left(-\frac{1}{2}\,\frac{(k - r(1-p)/p)^2}{r(1-p)/p^2}\right).$$

Полагая

$$\tau_r^{\bullet}=(\tau_r-r(1-p)/p)(r(1-p)/p^2)^{-1/2},$$

стандартными рассуждениями получаем интегральную предельную теорему

$$\mathbf{P}_{\tau_r^{\bullet}}(I_{a,b})\to\int_a^b\varphi(x)\,dx,\quad \varphi(x)=\left(1/\sqrt{2\pi}\right)e^{-x^2/2}.$$

$$\text{(II)}\ \mathbf{P}(\omega:\tau_n(\omega)>k)=\mathbf{P}(\omega:S_{k+r}(\omega)<r)=$$

$$=\mathbf{P}\left(\omega:\frac{S_{k+r}(\omega)-(k+r)p}{((k+r)p(1-p))^{1/2}}<\frac{r-(k+r)p}{((k+r)p(1-p))^{1/2}}\right)\to\Phi(x),$$

если $k+r\to\infty$, $(r(1-p)-kp)((k+r)p(1-p))^{-1/2}\to x$, $-\infty<x<\infty$, $\Phi(x)$ — функция распределения стандартного нормального закона. Последнее, очевидно, возможно, лишь когда k и r одного порядка, а следовательно, $r(1-p)-kp=O\left(\sqrt{k}\right)$. В частности, $r(1-p)\sim kp$, $k+r\sim r/p$ и

$$(r(1-p)/p-k)(r(1-p)/p^2)^{-1/2}\to x.$$

Таким образом,

$$\Phi(x)=\lim_{n\to\infty}\mathbf{P}\left(\omega:\frac{\tau_n(\omega)-r(1-p)/p}{(r(1-p)/p^2)^{1/2}}>-x\right),$$

или

$$\mathbf{P}_{\tau_r^{\bullet}}((-\infty,\ x])=1-\mathbf{P}_{\tau_r^{\bullet}}((x,\ \infty))\to1-\Phi(-x)=\Phi(x).$$

Подчеркнем, что $\mathbf{M}\tau_r=r(1-p)/p$, $\mathbf{D}\tau_r=r(1-p)/p^2$ (см. 7.7, 8.7) и что τ_r представляется как сумма r независимых одинаково геометрически распределенных случайных величин (см. 4.6),

$$\tau_r^{\bullet}=(\tau_r-\mathbf{M}\tau_r)/\sqrt{\mathbf{D}\tau_r}.]$$

12.7. Задача. Показать, что последовательность распределений, задаваемых гамма-плотностями

$$f_n(x)=(x^n/n!)\,e^{-x},\quad x>0,\quad f_n(x)=0,\quad x\leqslant0,\quad n=1,\ 2,\ \ldots,$$

после надлежащего центрирования и нормирования сходится к нормальному закону.

[Так как $(d/dx)\ln f_n(x)=n/x-1$, то $f_n(x)$ достигает максимума в точке $x=n$, убывая в стороны от нее. По формуле Стирлинга

$$f_n(n)\sim(2\pi n)^{-1/2},\quad n\to\infty.$$

Выбирая $\alpha_n=n$ в качестве центрирующей последовательности и оставляя нормирующую последовательность $\beta_n\to\infty$ пока не определенной, запишем

$$\ln(f_n(n+x\beta_n)/f_n(n)) = n\ln(1+x\beta_n/n) - x\beta_n =$$
$$= -x^2\beta_n^2/(2n) + O(t^3\beta_n^3/n^2).$$

Отсюда легко понять, что β_n^2 должно быть того же порядка, что n. Полагая $\beta_n = \sqrt{n}$, обнаруживаем, что предел — стандартное нормальное распределение.

Отметим, что сл. в. τ_n с плотностью $f_n(x)$ представима как сумма n независимых одинаково экспоненциально распределенных случайных величин (см. 4.19), $\mathbf{M}\tau_n = \mathbf{D}\tau_n = n$ (см. 7.27, 8.13) и $\overset{\bullet}{\tau}_n = (\tau_n - \mathbf{M}\tau_n)/\sqrt{\mathbf{D}\tau_n}$ имеет предельное стандартное нормальное распределение. Наконец, заметим, что гамма-распределение с плотностью $f_n(x)$ — предельное для отрицательных биномиальных с параметрами $n\to\infty$, $p_n\to 0$, $np_n\to 1$ (см. 3.14). Полученный выше результат связывает предельные теоремы 3.14 и 12.6 в том же самом отношении, в каком находится утверждение 12.5 для биномиального распределения.]

12.8. Пример (эмпирическая функция распределения и порядковые статистики). Пусть $X_1, X_2, \ldots, X_n$ — независимые одинаково распределенные случайные величины с непрерывной функцией распределения $F(x)$. В математической статистике рассматривается вопрос об оценивании неизвестной ф. р. $F(x)$ по результатам наблюдений за сл. в. $X_1, X_2, \ldots, X_n$. Фиксируя x, свяжем со сл. в. X_i, $i=1, \ldots, n$, схему Бернулли, принимая за «успех» в i-м испытании осуществление события $X_i(\omega)\leqslant x$. Частоту успехов в n испытаниях здесь принято обозначать

$$F_n(x) \equiv F_n(x;\ X_1,\ X_2,\ \ldots,\ X_n),$$

опуская зависимость от сл. в. X_i, $i=1, \ldots, n$. По закону больших чисел для схемы Бернулли для любого $\varepsilon>0$

$$\mathbf{P}(\omega : |F_n(x) - F(x)| > \varepsilon) \to 0,\ n\to\infty.$$

Более точные границы для отклонения $F_n(x)$ от $F(x)$ дает теорема Муавра—Лапласа

$$\mathbf{P}(\omega : |F_n(x) - F(x)| > t(F(x)(1-F(x))^{1/2} n^{-1/2}) \to 2(1-\Phi(t)),$$

где $\Phi(t)$ — функция распределения стандартного нормального закона. Эти соотношения позволяют считать $F_n(x)$, по крайней мере асимптотически, разумной оценкой для $F(x)$. Отсюда естественно предложить $F_n(x)$ как функцию x в качестве оценки для ф. р. $F(x)$, $-\infty<x<\infty$; $F_n(x)$ как функцию переменного x, зависящую от случайных параметров $X_1, \ldots, X_n$, называют *эмпирической функцией распределения* (э. ф. р.). Представив $F_n(x)$ в виде

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{\omega\,:\,X_i(\omega)\leqslant x\}}(\omega),$$

замечаем, что $F_n(x)$ — кусочно-постоянная неубывающая функция x со скачками величины $1/n$ в случайных точках X_i, $i=1, 2, \ldots$ $\ldots, n$ (либо кратными величине $1/n$, когда некоторые из сл. в. X_i, $i=1, \ldots, n$, приняли одинаковые значения). На рис. 33 реализация э. ф. р. $F_n(x)$ построена по выборке из равномерного распределения на отрезке $[0, 1]$.

Немаловажный интерес в математической статистике помимо оценки ф. р. $F(x)$ имеет оценка функции, обратной к $F(x)$. Положим для простоты, что $F(x)$ строго монотонна (в области, где $0<F(x)<1$) и обозначим для любого $0<p<1$ через x_p значение, при котором $F(x_p)=p$; x_p называют *р-квантилью* распределения F. Для оценки x_p было бы естественно предложить решение аналогичного уравнения для э. ф. р. $F_n(x)=p$. Однако поскольку $F_n(x)$ — ступенчатая, то решение этого уравнения не существует при p, не кратных $1/n$. Поэтому в качестве замены берут наименьшее x, удовлетворяющее неравенству $F_n(x)\geqslant p$. Пусть $X_{1n}\leqslant X_{2n}\leqslant\ldots\leqslant X_{nn}$ — порядковые статистики, образованные по набору сл. в. $X_1, X_2, \ldots, X_n$ ($X_{kn}(\omega)$ при заданном ω есть k-е в порядке возрастания из чисел $X_i(\omega)$, $i=1, \ldots, n$). Легко видеть, что указанное решение неравенства $F_n(x)\geqslant p$ есть $X_{[np], n}$ — порядковая статистика с номером, равным целой части np.

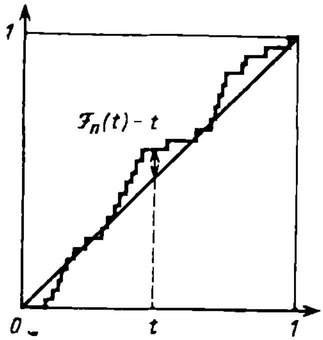

Рис. 33

Рассмотрим частный случай равномерного на отрезке $[0, 1]$ распределения: $F(x)=x$, $0\leqslant x\leqslant 1$, и установим предельную теорему для распределений сл. в. $X_{[np], n}$, $n\to\infty$, $0<p<1$, фиксировано (в общем случае см. 13.11).

Для доказательства предложим несколько вариантов. Можно исходить из формулы 3.4 для ф. р. сл. в. $X_{k, n}$:

$$F_{X_{k,n}}(x)=\sum_{j=k}^{n} C_n^j x^j (1-x)^{n-j}, \quad 0\leqslant x\leqslant 1,$$

либо из более компактной записи в терминах случайных величин (ср. 3.4)

$$\mathbf{P}(\omega : X_{k,n}(\omega)\leqslant x)=\mathbf{P}(\omega : nF_n(x)\geqslant k),$$

применяя к выражениям справа предельную теорему Муавра—Лапласа. По-видимому, более нагляден прямой метод исследования плотности (см. 3.10)

$$f_{X_{k,n}}(x)=nC_{n-1}^{k-1}x^{k-1}(1-x)^{n-k}=x^{-1}kp_n(k;\ x).$$

Поскольку

$$p_n(k;\ x)\sim(nx(1-x))^{-1/2}\varphi\left((k-nx)/\sqrt{nx(1-x)}\right)$$

при $k-nx=O(nx(1-x))^{1/2}$, то, выбирая x при $k=[np]$ равным $x=$ $=p+t(p(1-p)/n)^{1/2}$, имеем

$$f_{X_{k,n}}(x)\sim(p(1-p)/n)^{-1/2}\varphi(-t),\quad k=[np],$$

откуда видно, что сл. в. $X^*_{k,n}=\sqrt{n/(p(1-p))}(X_{k,n}-p)$ имеют в пределе стандартное нормальное распределение.

12.9.* Задача. Пусть $F_n(t)=F_n(t;\ X_1,\ \ldots,\ X_n)$ — эмпирическая функция распределения для последовательности $X_1,\ \ldots,\ X_n$ независимых случайных величин с равномерным распределением на отрезке $[0,\ 1]$. Вычислить совместное распределение вероятностей сл. в. $F_n(t_i)$, $i=1,\ \ldots,\ k$, $0<t_1<\ldots<t_k<1$. Показать, что конечномерные распределения процесса $X_n(t)=\sqrt{n}(F_n(t)-t)$, $0\leqslant t\leqslant 1$ (рис. 33), сходятся к распределениям броуновского моста (см. 11.10).

[Пусть $0\leqslant n_1\leqslant n_2\leqslant\ldots\leqslant n_k\leqslant n$ — произвольные целые числа. Если исключить событие вероятности нуль, что хотя бы один скачок случайного процесса $F_n(t)$, $0\leqslant t\leqslant 1$, пришелся на какую-либо из точек t_i, $i=1,\ \ldots,\ k$, то событие $\{nF_n(t_i)=n_i,\ i=1,\ \ldots,\ k\}$ происходит тогда и только тогда, когда в интервалы $(t_{i-1},\ t_i)$ попало ровно n_i-n_{i-1} значений из ряда сл. в. $X_1,\ \ldots,\ X_n$, $i=1,\ \ldots,\ k$ $(t_0=0,\ n_0=0)$. Поэтому (см. 1.25).

$$\mathbf{P}(\omega:nF_n(t_i)=n_i,\ i=1,\ \ldots,\ k)=$$
$$=n!\,(n_1!\,(n_2-n_1)!\,\ldots(n_k-n_{k-1})!\,(n-n_k)!)^{-1}\times$$
$$\times t_1^{n_1}(t_2-t_1)^{n_2-n_1}\ldots(t_k-t_{k-1})^{n_k-n_{k-1}}(1-t_k)^{n-n_k},$$

откуда находим

$$p_{nF_n(t_k)|nF_n(t_1),\ldots,nF_n(t_{k-1})}(n_k\,|\,n_1,\ \ldots,\ n_{k-1})=$$
$$=\frac{(n-n_{k-1})!}{(n_k-n_{k-1})!\,(n-n_k)!}\cdot\frac{(t_k-t_{k-1})^{n_k-n_{k-1}}(1-t_k)^{n-n_k}}{(1-t_{k-1})^{n-n_{k-1}}}\equiv p(t_{k-1},\ n_{k-1},\ t_k,\ n_k).$$

Таким образом, последовательность сл. в. $nF_n(t_1),\ \ldots,\ nF_n(t_k)$ образует марковскую цепь при любых k, $t_1<\ldots<t_k$, а сам процесс $nF_n(t)$ ввиду этого называется *марковским с переходной функцией* $p(t,\ m,\ s,\ l)$, $t<s$.

Положим $x_i=\sqrt{n}(n_i/n-t_i)$, $i=1,\ 2$, $0<t_1<t_2<1$. Заметив, что

$$p(t_1,\ n_1,\ t_2,\ n_2)=p_{n-n_1}(n_2-n_1;\ (t_2-t_1)/(1-t_1)),\quad t_1<t_2,$$

где справа стоит биномиальное распределение, имеем по теореме Муавра—Лапласа

$$p_{X_n(t_2)\mid X_n(t_1)}(x_2\mid x_1)\sim\left(2\pi(n-n_1)\frac{t_2-t_1}{1-t_1}\cdot\frac{1-t_2}{1-t_1}\right)^{-1/2}\times$$

$$\times\exp\left(-\frac{1}{2}\left(n_2-n_1-(n-n_1)\frac{t_2-t_1}{1-t_1}\right)^2\Big/\left((n-n_1)\frac{t_2-t_1}{1-t_1}\cdot\frac{1-t_2}{1-t_1}\right)\right)\sim$$

$$\sim(2\pi n)^{-1/2}\left(\frac{1-t_1}{(t_2-t_1)(1-t_2)}\right)^{1/2}\exp\left(-\frac{1}{2}\,\frac{(x_2-x_1(1-t_2)/(1-t_1))^2}{(t_2-t_1)(1-t_2)/(1-t_1)}\right).$$

Сравнивая полученное выражение с асимптотикой для переходных вероятностей условного случайного блуждания в 11.10 и замечая, что

$$\frac{\varphi(y/\sqrt{1-t})}{\varphi(x/\sqrt{1-s})}\,\varphi\left((y-x)/\sqrt{t-s}\right)=\exp\left(-\frac{1}{2}\left(\frac{(y-x)^2}{t-s}+\frac{y^2}{1-t}-\frac{x^2}{1-s}\right)\right)=$$

$$=\exp\left(-\frac{1}{2}(y-x(1-t)/(1-s))^2/((t-s)(1-t)/(1-s))\right),$$

получаем требуемый результат стандартным переходом к интегральной теореме.]

12.10. Пример (статистики Колмогорова—Смирнова). Одна из постановок задач математической статистики состоит в следующем. Предполагают, что наблюдения $x_1, x_2, \ldots, x_n$ получаются как реализация последовательности $X_1, X_2, \ldots, X_n$ независимых одинаково распределенных случайных величин с некоторой непрерывной функцией распределения. Высказывается *гипотеза*, что эта ф. р. есть заданная $F(x)$. Требуется проверить, согласуются ли результаты наблюдения с этой гипотезой. Представляется естественным рассмотреть для этого какую-либо меру отклонения э. ф. р. $F_n(x)$ от заданной ф. р. $F(x)$, $-\infty<x<\infty$. Наиболее простыми мерами можно считать статистики Колмогорова $D_n=\sup_{\{x\}}|F_n(x)-F(x)|$ и Смирнова $D_n^+=\sup_{\{x\}}(F_n(x)-F(x))$, $D_n^-=\inf_{\{x\}}(F_n(x)-F(x))$. Статистика D_n более универсальна, статистики D_n^+ и D_n^- употребляются, когда есть основания предполагать, что при нарушении гипотезы отклонение истинной ф. р. будет происходить в сторону больших и соответственно меньших ее значений. Замечательно, что распределения вероятностей D_n, D_n^+, D_n^- не зависят от $F(x)$, что дает возможность рассчитать их для какой-то наиболее удобной для вычислений ф. р. Проверим сделанное утверждение.

Заметим прежде всего, что если сл. в. X имеет непрерывную ф. р. $F(x)$, то сл. в. $Y=F(X)$ равномерно распределена на отрезке $[0, 1]$:

$$\mathbf{P}(\omega: Y(\omega)\leqslant y)=\mathbf{P}(\omega: X(\omega)\leqslant F^{-1}(y))=F(F^{-1}(y))=y,$$

$$0<y<1,\quad F^{-1}(y)=\sup\{x: F(x)=y\}.$$

Таким образом, сл. в. $Y_i=F(X_i)$, $i=1, \ldots, n$, независимы и равномерно распределены на отрезке $[0, 1]$. Далее имеем при $0<y<1$

$$F_n(y;\ Y_1,\ \ldots,\ Y_n)=\frac{1}{n}\sum_{i=1}^{n} I_{\{\omega: Y_i(\omega)\leqslant y\}}(\omega)=$$

$$=\frac{1}{n}\sum_{i=1}^{n} I_{\{\omega: X_i(\omega)\leqslant F^{-1}(y)\}}(\omega)=F_n(F^{-1}(y);\ X_1,\ \ldots,\ X_n),$$

т. е. процесс $F_n(x;\ X_1, \ldots, X_n)$, $-\infty<x<\infty$, получается из процесса $F_n(y;\ Y_1,\ \ldots,\ Y_n)$, $0<y<1$, монотонным преобразованием параметра $y=F(x)$. Так как этим же преобразованием связаны ф. р. сл. в. X и Y, а статистики D_n, D_n^+, D_n^- измеряют расстояние по вертикали, то их распределения не зависят от преобразования оси абсцисс. Например,

$$\sup_{\{x: 0<F(x)<1\}} |F_n(x;\ X_1,\ \ldots,\ X_n)-F(x)|=$$

$$=\sup_{\{y: 0<y<1\}} |F_n(F^{-1}(y));\ X_1,\ \ldots,\ X_n)-F(F^{-1}(y))|=$$

$$=\sup_{\{y: 0<y<1\}} |F_n(y;\ Y_1,\ \ldots,\ Y_n)-y|,$$

и остается заметить, что на множестве тех x, где $F(x)=0$ или $F(x)=1$, с вероятностью 1 $F_n(x;\ X_1,\ \ldots,\ X_n)=F(x)$.

Пусть теперь X_i, $i=1,\ \ldots,\ n$, распределены равномерно на $[0, 1]$. В 12.9 показано, что процесс $X_n(t)=\sqrt{n}\,(F_n(t)-t)$, $0\leqslant t\leqslant 1$, сходится по распределению к броуновскому мосту и то же имеет место для процесса $(S_{[nt]}\,|\,\Lambda_n)/\sqrt{n}$ (см. 11.10). Но (см. 11.12)

$$\mathbf{P}\left(\omega: \max_{\{t\}}(S_{[nt]}\,|\,\Lambda_n)/\sqrt{n}\leqslant x\right)\to 1-e^{-2x^2},\quad n\to\infty.$$

Есть основания надеяться, что также

$$\mathbf{P}(\omega: \max_{\{t\}} X_n(t)\leqslant x)\to 1-e^{-2x^2},\quad n\to\infty.$$

Доказательство этого факта выходит за рамки данной книги, однако действительно для функционалов, непрерывно зависящих от траектории $x(t)$ случайного процесса в равномерной метрике $\sup_{\{t\}}|x(t)-y(t)|$, такого рода утверждения справедливы.

12.11. Пример (усиленный закон больших чисел, лемма Бореля—Кантелли). Эмпирический закон устойчивости частот утверждает, что в неограниченно продолжающейся последовательности повторных испытаний частота появления определенного события стремится к некоторому пределу. Вероятностная модель появлений заданного события при повторных наблюдениях есть испытания Бернулли, а соответствующая эмпирической закономерности математическая теорема могла бы состоять в утверждении, что $n^{-1}S_n(\omega)\to p$ при $n\to\infty$, где $S_n(\omega)$ — число успехов в первых n испытаниях. Вероятностная модель последовательности повторных испытаний допускает любую последовательность из 0 и 1 в качестве возможных исходов испытаний Бернулли, продолжающихся неограниченно, но за исключением некоторого множества вероятности нуль, действительно $n^{-1}S_n(\omega)\to p$; при этом говорят, что с вероятностью 1 последовательность сл. в. $n^{-1}S_n(\omega)$ сходится к p, а само утверждение называется *усиленным законом больших чисел* (УЗБЧ) для схемы Бернулли. Докажем его.

Подмножество A тех $\omega\in\Omega$, для которых последовательность функций $n^{-1}S_n(\omega)$ сходится к p, можно представить в виде

$$A=\bigcap_{\{k\}}\bigcup_{\{N\}}\bigcap_{\{n>N\}}\{\omega:|n^{-1}S_n(\omega)-p|<2^{-k}\}.$$

В самом деле, если при $\omega=\omega_0$ $n^{-1}S_n(\omega_0)\to p$ при $n\to\infty$, то для любого $\varepsilon=2^{-k}$, k — натуральное, найдется номер $N=N_k$, такой, что при всех $n>N_k$ имеет место неравенство $|n^{-1}S_n(\omega_0)-p|<2^{-k}$, это и означает $\omega_0\in A$. Точно так же из $\omega_0\in A$ следует $n^{-1}S_n(\omega)\to p$, $n\to\infty$.

Докажем, что $\mathbf{P}(\bar{A})=0$ и положим для краткости $A_{n,k}=\{\omega:n^{-1}|S_n(\omega)-p|<2^{-k}\}$. Тогда

$$\bar{A}=\bigcup_{\{k\}}\bigcap_{\{N\}}\bigcup_{\{n>N\}}\bar{A}_{n,k}$$

и $\mathbf{P}(\bar{A})=0$ тогда и только тогда, когда при любом k

$$\mathbf{P}\Big(\bigcap_{\{N\}}\bigcup_{\{n>N\}}\bar{A}_{n,k}\Big)=0.$$

[В одну сторону последнее утверждение очевидно, в другую оно является следствием свойства *полуаддитивности* вероятностной меры:

$$\mathbf{P}\Big(\bigcup_{i=1}^{\infty}A_i\Big)\leqslant\sum_{i=1}^{\infty}\mathbf{P}(A_i)$$

для любых событий A_i, $i=1, 2, \ldots$, а оно в свою очередь вытекает из свойства непрерывности вероятностной меры (см. § 6). Именно, полагая $B_n=\bigcup_{i=1}^{n}A_i$, получаем

$$\lim_{n\to\infty}\mathbf{P}(B_n)=\mathbf{P}\Big(\bigcup_{i=1}^{\infty}A_i\Big),$$

в то время как

$$\mathbf{P}(B_n)=\mathbf{P}\Big(\bigcup_{i=1}^{n}A_i\Big)\leqslant\sum_{i=1}^{n}\mathbf{P}(A_i).]$$

Далее заметим, что последовательность событий

$$\bigcup_{\{n>N\}}\bar{A}_{n,k},\ N=1, 2, \ldots,$$

— убывающая, и по свойству непрерывности вероятностной меры

$$\mathbf{P}\Big(\bigcap_{\{N\}}\bigcup_{\{n>N\}}\bar{A}_{n,k}\Big)=\lim_{N\to\infty}\mathbf{P}\Big(\bigcup_{\{n>N\}}\bar{A}_{n,k}\Big).$$

Равенство нулю предела справа получаем, воспользовавшись оценкой

$$\mathbf{P}\Big(\bigcup_{\{n>N\}}\bar{A}_{n,k}\Big)\leqslant\sum_{\{n>N\}}\mathbf{P}(\bar{A}_{n,k})$$

и заметив, что ряд из $\mathbf{P}(\bar{A}_{n,k})$ сходится, так как вероятности

$$\mathbf{P}(\bar{A}_{n,k})=\mathbf{P}(\omega:S_n(\omega)\geqslant n(p+2^{-k}))+\mathbf{P}(\omega:S_n(\omega)\leqslant n(p-2^{-k}))$$

экспоненциально убывают с ростом n при любом k (см. 12.4). Тем самым УЗБЧ установлен. Выделим из его доказательства одно часто используемое утверждение.

Лемма Бореля—Кантелли: *если последовательность событий* A_n, $n=1, 2, \ldots$, *такова, что* $\sum_{n=1}^{\infty}\mathbf{P}(A_n)<\infty$, *то с вероятностью* 1 *осуществляется лишь конечное число событий из этой последовательности.*

Событие, на котором осуществляется бесконечно много событий, можно записать в виде

$$\bigcap_{\{N\}}\bigcup_{\{n>N\}}A_n,$$

а оно, как мы видели в случае $A_n=\bar{A}_{n,k}$, имеет вероятность нуль.

12.12. Задача. Пользуясь леммой Бореля—Кантелли, показать, что с вероятностью 1 происходит лишь конечное число событий последовательности $A_n=\{\omega:|S_n(\omega)-np|>((2+\varepsilon)np(1-p)\times\times\ln n)^{1/2}$, $n=1, 2, \ldots$, где S_n — число успехов в n испытаниях Бернулли, $\varepsilon>0$.

[Используя оценку больших уклонений из 12.1, получаем

$$\mathbf{P}(A_n)\leqslant(2\pi(2+\varepsilon)\ln n)^{-1/2}n^{-1-\varepsilon/2},$$

откуда результат следует применением леммы Бореля—Кантелли. Отсюда видно, что $n^{-1}S_n(\omega)-p=O(1)(\ln n/n)^{1/2}$, правда $O(1)$ есть функция от ω.]

12.13. Задача. Доказать *вторую лемму Бореля—Кантелли*: если события A_1, A_2, ... взаимно независимы и ряд, составленный из вероятностей $\mathbf{P}(A_n)$, расходится, то с вероятностью единица осуществляется бесконечно много событий последовательности A_n, $n=1, 2, \ldots$.

[Событие, состоящее в осуществлении конечного числа событий последовательности A_n, $n=1, 2, \ldots$, представимо в виде

$$\bigcup_{\{N\}}\bigcap_{\{n>N\}}\bar{A}_n.$$

Оно имеет вероятность нуль тогда и только тогда, когда

$$\mathbf{P}\Big(\bigcap_{\{n>N\}}\bar{A}_n\Big)=0$$

(необходимость очевидна, достаточность вытекает из полуаддитивности вероятностной меры). Последовательность событий $\bigcap\limits_{=N+1}^{N+k}\bar{A}_N$, $k=1, 2, \ldots,$ возрастает, а потому

$$\mathbf{P}\Big(\bigcap_{\{n>N\}}\bar{A}_n\Big)=\lim_{k\to\infty}\mathbf{P}\Big(\bigcap_{n=N+1}^{N+k}\bar{A}_n\Big)=$$

$$=\lim_{k\to\infty}\prod_{n=N+1}^{N+k}\mathbf{P}(\bar{A}_n)=\prod_{n=N+1}^{\infty}(1-\mathbf{P}(A_n)).$$

Но произведение обращается в нуль тогда и только тогда, когда расходится ряд из $\mathbf{P}(A_n)$.]

12.14. Задача. Пусть последовательность сл. в. $X_n(\omega)$ сходится с вероятностью 1 к сл. в. $X(\omega)$: $\mathbf{P}(\omega:\lim\limits_{n\to\infty}X_n(\omega)=X(\omega))=1$. Показать, что для любого $\varepsilon>0$

$$\mathbf{P}(\omega:|X_n(\omega)-X(\omega)|>\varepsilon)\to 0,\quad n\to\infty.$$

[Множество сходимости A последовательности сл. в. $X_n(\omega)$ к конечному пределу $X(\omega)$ можно представить в виде

$$A=\bigcap_{\{k\}}\bigcup_{\{N\}}\bigcap_{\{n>N\}}\{\omega:|X_n(\omega)-X(\omega)|<2^{-k}\}.$$

Рассуждения по схеме доказательства УЗБЧ в 12.11 показывают, что из $\mathbf{P}(\bar{A})=0$ следует при любом k

$$\lim_{N\to\infty}\mathbf{P}\Big(\bigcup_{\{n>N\}}\{\omega:|X_n(\omega)-X(\omega)|\geqslant 2^{-k}\}\Big)=0,$$

а следовательно, и

$$\lim_{n\to\infty}\mathbf{P}(\omega:|X_n(\omega)-X(\omega)|\geqslant 2^{-k})=0.]$$

Далее мы рассматриваем простое случайное блуждание $S_n=X_1+\ldots+X_n$, $n=1, 2, \ldots,$ $S_0=0$, где X_i, $i=1, 2, \ldots,$ — независимые случайные величины, принимающие значения 1 и -1 с вероятностями p и $1-p$ соответственно. Поскольку сл. в. $Y_i=(X_i+1)/2$, $i=1, 2, \ldots,$ образуют последовательность испытаний Бернулли, то полученные выше результаты для схемы Бернулли позволяют заключить, например, что

$$\mathbf{P}(\omega:\lim_{n\to\infty}n^{-1}S_n(\omega)=2p-1)=1.$$

Отсюда, кстати, следует, что с вероятностью 1 последовательность $S_n(\omega)$, $n=1, 2, \ldots,$ стремится к ∞ при $p>1/2$ и к $-\infty$ при

$p<1/2$ (другое доказательство см. 12.15). Из теоремы Муавра—Лапласа имеем

$$\mathbf{P}(\omega : (S_n(\omega)-n(2p-1))/\sqrt{4np(1-p)}\leqslant x)\to\Phi(x),\quad n\to\infty.$$

Отметим, что $\mathbf{M}S_n=n(2p-1)$, $\mathbf{D}S_n=4np(1-p)$.

12.15.* Задача. Обозначим $T_m=T_m(\omega)$ момент первого достижения случайным блужданием $S_n(\omega)$, $n=1, 2, \ldots$, положения m, считая $T_m(\omega)=\infty$ для тех ω, при которых $S_n(\omega)\neq m$ ни при каком m. Вывести из 2.20, что при $p\geqslant 1/2$ и $m>0$ $\mathbf{P}(\omega : T_m(\omega)<<\infty)=1$.

[Событие $A=\{\omega : T_m(\omega)<\infty\}$, $m\geqslant 1$, представим в виде объединения

$$A=\bigcup_{N=1}^{\infty}\{\omega : T_m(\omega)<T_{-N}(\omega)\}$$

возрастающей последовательности событий $A_N=\{\omega : T_m(\omega)<<T_{-N}(\omega)\}$, $N=1, 2, \ldots$. По свойству непрерывности вероятности

$$\mathbf{P}(A)=\lim_{N\to\infty}\mathbf{P}(A_N).$$

Однако по формуле вероятности разорения игрока 2.20 (при $b=N+m$, $a=N$) имеем

$$\mathbf{P}(\overline{A}_N)=\frac{\lambda^N+\lambda^{N+1}+\ldots+\lambda^{N+m-1}}{1+\lambda+\ldots+\lambda^{N+m-1}},\quad \lambda=(1-p)/p,$$

что, очевидно, стремится к нулю при $N\to\infty$ и фиксированном m в предположении $p\geqslant 1/2$, т. е. $\lambda=(1-p)/p\leqslant 1$.]

12.16. Задача. Показать, что конечномерные распределения последовательности процессов

$$X_n(t)=(S_{[nt]}-tn(2p-1))/\sqrt{4np(1-p)},\quad 0\leqslant t\leqslant 1,\ n=1, 2, \ldots,$$

где S_k, $k=1, 2, \ldots$, — простое случайное блуждание, сходятся к распределениям броуновского движения.

[Проверяется аналогично симметричному случаю $p=1/2$ (см. § 11).]

12.17. Пример (максимум блуждания и времена достижения). Нетрудно видеть, что сл. в. T_m — момент первого достижения простым случайным блужданием положения m, как и в симметричном случае, является моментом остановки. Однако здесь $T_m(\omega)$ может принимать бесконечное значение с положительной вероятностью, поскольку при $p\neq 1/2$ блуждание с вероятностью 1 уходит в бесконечность того же знака, что и число $2p-1$. Пусть для определенности $p>1/2$. Тогда $\mathbf{P}(\omega : T_k(\omega)<\infty)=1$, $k==1, 2, \ldots$. Распределение вероятностей сл. в. T_k легко вывести из комбинаторных формул для симметричного блуждания. Все пути длины n, для которых $T_k(\omega)=n$, имеют равные вероятности $p^{(n+k)/2}\cdot(1-p)^{(n-k)/2}$, так что из 11.7 получаем

$$\mathbf{P}(\omega : T_k(\omega)=n)=(k/n)\,C_n^{(n+k)/2}p^{(n+k)/2}(1-p)^{(n-k)/2}$$

Предельное распределение для последовательности сл. в. $T_k(\omega)$, $k=1, 2, \ldots$, может быть получено теми же приемами, что и в 12.5—12.7 (см. 12.18). Предельное распределение для максимума блуждания $M_n=\max(S_0, S_1, \ldots, S_n)$ также получается из стандартного соображения: $T_k>n \Leftrightarrow M_n<k$ (см. 12.9). Здесь рассмотрим предельное распределение пары сл. в. M_n, S_n. Их совместное распределение содержится в 2.13 и равно

$$p_{M_n,S_n}(r, k)=\frac{2r-k+1}{n+1}\,C_{n+1}^{r+1+(n-k)/2}p^{(n+k)/2}(1-p)^{(n-k)/2}$$

при $r\geqslant k\geqslant -n+2r$ ($C_m^s=0$ для $s<0$, $s>m$ и s не целых).

Для получения предельной теоремы для пары (M_n, S_n) необходимо понять, где расположена основная масса распределения. Проведем это исследование, опираясь на вероятностную интуицию, а затем подтвердим все вычислениями. Легко понять (рис. 34), что максимум блуждания достигается в такой момент T_k, что $T_k\leqslant n<T_{k+1}$, при этом $M_n=T_k$. В соответствии с 11.3, 11.4 участки блуждания между случайными моментами T_i, T_{i+1} независимы и с точностью до параллельного переноса одинаково распределены.

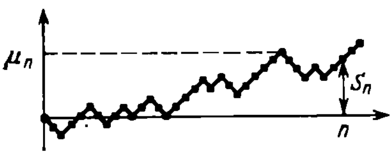

Рис. 34

Каждый из интервалов (T_i, T_{i+1}) имеет такое же распределение, как и T_1, а потому можно предположить, что интервал, накрывающий точку n, не будет разрастаться с ростом n (не следует думать, что он имеет то же распределение, что и T_1, см. по этому поводу 3.13). Но тогда (рис. 34) и разность между M_n и S_n должна быть конечной, а не растущей с ростом n. Поскольку предельное распределение для S_n нам уже известно, то есть основания полагать, что основная масса распределения вероятностей $p_{M_n,S_n}(r, k)$ расположена в точках (r, k) вида $k=n(2p-1)+ +x(4np(1-p))^{1/2}$, $r=k+m$, где x с ростом n меняется в разрастающейся, но не слишком быстро, области. Асимптотику для $p_{M_n,S_n}(r, k)$ достаточно искать при ограниченных x и m.

При указанных r и k имеем при $n\to\infty$

$$p_{M_n,S_n}(r, k)\sim(2p-1)\,C_{n+1}^{(n-k)/2+r+1}p^{(n+k)/2}(1-p)^{(n-k)/2}=$$

$$=(2p-1)\,p_{n+1}((n+k)/2+m+1;\ p)\,p^{-m-1}(1-p)^m,$$

где $p_n(k;\ p)$ — биномиальное распределение. Применяя теорему Муавра—Лапласа, получаем

$$p_{M_n, S_n}(r,\ k) \sim (np(1-p))^{-1/2}\,\varphi(x)\,((2p-1)/p)\,((1-p)/p)^m.$$

С учетом равенства

$$p_{M_n-S_n, S_n}(m,\ k) = p_{M_n, S_n}(r,\ k), \quad r = k+m,$$

суммируя по k, находим при $n \to \infty$

$$\mathrm{P}(\omega : M_n(\omega) - S_n(\omega) = m,\ \ (S_n(\omega) - n(2p-1))\,(4np(1-p))^{-1/2} \leqslant x) \to$$
$$\to \Phi(x)\,((2p-1)/p)\,((1-p)/p)^m.$$

Отсюда заключаем, что

$$\mathrm{P}(\omega : M_n(\omega) - S_n(\omega) = m) \to ((2p-1)/p)\,((1-p)/p)^m, \quad m = 0,\ 1,\ 2, \ldots .$$

Так как совместное предельное распределение сл. в. $M_n - S_n$ и S_n распадается в произведение частных предельных распределений $M_n - S_n$ и S_n, то говорят, что сл. в. *асимптотически независимы*. Наконец,

$$\frac{M_n - n(2p-1)}{(4np(1-p))^{1/2}} = \frac{S_n - n(2p-1)}{(4np(1-p))^{1/2}} + \frac{M_n - S_n}{(4np(1-p))^{1/2}},$$

откуда нетрудно вывести (см. § 13), что сл. в. $(M_n - n(2p-1)) \times (4np(1-p))^{-1/2}$ имеет стандартное нормальное предельное распределение.

12.18. Задача. Вывести предельную теорему для последовательности времен достижения T_k, $k = 1,\ 2,\ \ldots$, при $p > 1/2$, используя явную формулу для распределения T_k из 12.17.

[Отношение

$$p_{T_k}(n+2)/p_{T_k}(n) = 4n(n+1)\,p(1-p)/((n+2)^2 - k^2)$$

переходит единицу вблизи $n = [k/(2p-1)]$. Предположим, что $(n+k)/2 - np = O(\sqrt{n})$, т. е. $n - k/(2p-1) = O(\sqrt{k})$, где $n,\ k \to \infty$, и применим теорему Муавра—Лапласа:

$$p_{T_k}(n) \sim (2p-1)\,(np(1-p))^{-1/2}\,\varphi\left(((n+k)/2 - np)/\sqrt{np(1-p)}\right) \sim$$
$$\sim (kp(1-p)(2p-1)^{-3})^{-1/2}\,\varphi\left(\frac{-n + k/(2p-1)}{2\,(p(1-p)/(2p-1)^3)^{1/2}}\right), \quad n \to \infty.$$

Полагая $n - k/(2p-1) = x\,(p(1-p)\,(2p-1)^{-3})^{1/2}$, получаем

$$p_{T_k}(n) \sim (kp(1-p)(2p-1)^{-3})^{-1/2}\,\varphi(x),$$

что дает стандартный нормальный закон для предельного распределения сл. в. $(T_k - k/(2p-1)) \cdot (k \cdot 2p(1-p)\,(2p-1)^{-3})^{-1/2}$.]

12.19. Задача. Вывести из 12.18 предельную теорему для максимума случайного блуждания M_n (см. 12.17).

[Используя равенство $\{\omega : T_k(\omega) < n\} = \{\omega : M_n(\omega) > k\}$, имеем

$$\mathbf{P}\left(\omega : \frac{T_k(\omega) - ak}{\sigma\sqrt{k}} < \frac{n - ak}{\sigma\sqrt{k}}\right) = \mathbf{P}\left(\omega : \frac{M_n(\omega) - n/a}{\sigma\sqrt{k/a}} > \frac{ka - n}{\sigma\sqrt{k}}\right).$$

Пусть $n \to \infty$, $(n - ak)/(\sigma\sqrt{k}) \to x$, $a = (2p-1)^{-1}$, $\sigma = 2\left(p(1-p) \times \times (2p-1)^{-3}\right)^{1/2}$. Тогда левая часть стремится к $\Phi(x)$. Учитывая, что при этом $n \sim ak$, получаем

$$\mathbf{P}(\omega : (M_n(\omega) - n/a)/(\sigma n^{1/2} a^{-3/2}) > -x) \to \Phi(x).$$

Это совпадает с результатом, анонсированным в 12.17.]

12.20. Задача. Найти асимптотику хвоста распределения сл. в. T_1 при $p > 1/2$.

$$[\mathbf{P}(\omega : T_1(\omega) \geqslant n) = \sum_{\{k \geqslant n\}} k^{-1} C_k^{(k+1)/2} p^{(k+1)/2} (1-p)^{(k-1)/2},$$

причем отношение соседних ненулевых членов суммы справа равно (см. 12.18)

$$p_{T_1}(k+2)/p_{T_1}(k) = 4k(k+1)\,p(1-p)/((k+2)^2 - 1) \sim 4p(1-p) < 1$$

при $k \to \infty$. Это означает, что, вынося за знак суммы первый член в качестве общего множителя, можно заменить члены остающейся суммы геометрической прогрессией со знаменателем $4p(1-p)$ и получить при нечетном n

$$\mathbf{P}(\omega : T_1(\omega) \geqslant n) \sim n^{-1} C_n^{(n+1)/2} p^{(n+1)/2} (1-p)^{(n-1)/2} \cdot (1-2p)^{-2} \sim$$
$$\sim (2\pi)^{1/2} n^{-3/2} (p(1-p))^{n/2}\, 2^n p^{1/2} (1-p)^{-1/2} (1-2p)^{-2}.]$$

12.21. Пример (среднее, дисперсия и п. ф. времен достижения). Случайное блуждание можно начать не только из любой фиксированной точки $a : S_n^{(a)} = a + S_n$, $n = 0, 1, 2, \ldots$, но и из случайной: $S_n^{(Z)} = Z + S_n$, $n = 0, 1, 2, \ldots$, $S_0 = 0$, полагая, что $Z = Z(\omega)$ есть некоторая целочисленная случайная величина, не зависящая от последовательности сл. в. X_i, $i = 1, 2, \ldots$, определяющей дальнейшее движение частицы. Вычисление вероятностей событий, связанных с блужданием $S_n^{(Z)}$, $n = 0, 1, 2, \ldots$, сводится к применению формулы полной вероятности по полной группе событий $Z = m$, $m = 0, \pm 1, \pm 2, \ldots$. Случайные блуждания типа $S_n^{(Z)}$, $n = 0, 1, 2, \ldots$, возникают прежде всего в связи с моментами остановки. Как мы видели в 11.3, для любого момента остановки T простого случайного блуждания S_n, $n = 0, 1, 2, \ldots$; $S_0 = 0$, последовательность S_{n+T}, $n = 0, 1, 2, \ldots$, образует случайное блуждание со случайной начальной точкой S_T, причем это блуждание зависит от начального отрезка $S_0, S_1, \ldots, S_T$ только через значение S_T. Самый простой момент остановки $T = \text{const}$ оказывается чрезвычайно полезным при изучении простого случайного блуждания S_n, $n=0, 1, 2, \ldots, S_0=0$. Применим указанные соображения для вычисления некоторых характеристик этого блуждания.

Предположим для определенности, что $p>1/2$. Тогда сл. в. T_1 — момент первого достижения положения 1 — конечна с вероятностью 1 и, как следует из 12.20 и 7.3, имеет конечное математическое ожидание. Найдем его, вычисляя условное математическое ожидание T_1 относительно сл. в. S_1 (см. 9.8):

$$\mathbf{M}T_1=\mathbf{MM}(T_1|S_1).$$

Случайное блуждание S_{1+n}, $n=0, 1, 2, \ldots$, начинается из случайной точки S_1. Обозначим T_1^* время до первого попадания этого блуждания в положение 1, включая его начальное положение, так что $T_1^*=T_1-1$. Таким образом, $\mathbf{M}(T_1|S_1)=1+\mathbf{M}(T_1^*|S_1)$. Но $\mathbf{M}(T_1^*|\{S_1=1\})=0$, а при условии $S_1=-1$ условное распределение сл. в. T_1^* такое же, как безусловное распределение сл. в. T_2 — момента первого достижения положения 2 случайным блужданием S_n, $n=0, 1, 2, \ldots, S_0=0$. Поэтому

$$\mathbf{M}(T_1^*|\{S_1=-1\})=\mathbf{M}T_2=2\mathbf{M}T_1,$$

где мы воспользовались тем, что сл. в. T_2 представима в виде суммы двух (независимых) случайных величин, распределенных так же, как T_1. Итак,

$$\mathbf{M}T_1=1\cdot p+(1+2\mathbf{M}T_1)(1-p) \text{ или } \mathbf{M}T_1=1/(2p-1).$$

Аналогично находим дисперсию

$$\mathbf{D}T_1=\mathbf{D}T_1^*=\mathbf{MM}((T_1^*-\mathbf{M}T_1^*)^2|S_1)=p\cdot(\mathbf{M}T_1^*)^2+$$

$$+(1-p)\mathbf{M}((T_1^*-\mathbf{M}T_1^*)^2|\{S_1=-1\})=p(\mathbf{M}T_1-1)^2+$$

$$+(1-p)\mathbf{M}(T_2-\mathbf{M}T_1+1)^2,$$

$$\mathbf{M}(T_2-\mathbf{M}T_1+1)^2=\mathbf{M}(T_2-2\mathbf{M}T_1+\mathbf{M}T_1+1)^2=\mathbf{M}(T_2-2\mathbf{M}T_1)^2+$$

$$+(\mathbf{M}T_1+1)^2=\mathbf{D}T_2+(\mathbf{M}T_1+1)^2=2\mathbf{D}T_1+(\mathbf{M}T_1+1)^2,$$

$$\mathbf{D}T_1=p(\mathbf{M}T_1-1)^2+(1-p)2\mathbf{D}T_1+(1-p)(\mathbf{M}T_1+1)^2,$$

$$(2p-1)\mathbf{D}T_1=\mathbf{M}T_1^2+2\mathbf{M}T_1(1-2p)+1=4p(1-p)/(2p-1),$$

$$\mathbf{D}T_1=4p(1-p)/(2p-1)^2.$$

Из полученных формул вытекает, что в предельной теореме для T_k, $k=1, 2, \ldots$, в 12.18 в качестве центрирующих и нормирующих последовательностей берутся $\mathbf{M}T_k=k\mathbf{M}T_1$, $\mathbf{D}T_k=k\mathbf{D}T_1$, $k=1, 2, \ldots$.

Тот же подход можно применить к вычислению производящей функции (п. ф.) сл. в. T_1:

$$\mathcal{P}_{T_1}(t)=\mathbf{M}t^{T_1}=\mathbf{MM}(t^{T_1}|S_1)=\mathbf{MM}(t^{T_1^*+1}|S_1)=$$

$$=pt+(1-p)t\mathbf{M}(t^{T_1^*}|\{S_1=-1\})=pt+(1-p)t\mathbf{M}t^{T_2}.$$

Принимая во внимание, что сл. в. T_2 представима в виде суммы двух независимых величин, распределенных как T_1, получаем (см. 7.14), что $\mathcal{P}_{T_2}(t)=\mathcal{P}_{T_1}(t)^2$ и для нахождения п. ф. $\mathcal{P}_{T_1}(t)$ имеем квадратное уравнение

$$(1-p)\,t\mathcal{P}_{T_1}(t)^2-\mathcal{P}_{T_1}(t)+pt=0.$$

Оно имеет два корня, из которых один — неограниченная функция переменного t при $t\to 0$, а второй равен

$$(1-(1-4p(1-p)t^2)^{1/2})/(2(1-p)t).$$

Очевидно, что искомая п. ф. $\mathcal{P}_{T_1}(t)$, $0\leqslant t\leqslant 1$, при всех t выражается именно этой формулой. Распределение вероятностей сл. в. T_1 может быть получено разложением в ряд Тейлора (бином Ньютона с показателем 1/2). Так как $\mathcal{P}_{T_k}(t)=(\mathcal{P}_{T_1}(t))^k$, то и распределение сл. в. T_k легко получить разложением в ряд Тейлора.

12.22. Задача. Показать, что при $p<1/2$ условный процесс

$$X_n(t)=n^{-1/2}(S_{[nt]}\,|\,\Lambda_n),\quad \Lambda_n=\{\omega : S_n>0\},$$

где n — четно, S_k, $k=0, 1, 2, \ldots$, — простое случайное блуждание, $S_0=0$, сходится по распределению к броуновскому мосту (см. 11.10). Это же справедливо, если взять $\Lambda_n=\{\omega : S_n=0\}$.

[Общая схема рассуждений стандартна, выпишем лишь асимптотику для переходных вероятностей марковской цепи $Y_k=(S_k\,|\,\Lambda_n)$, $k=0, 1, \ldots, n$, для $\Lambda_n=\{\omega : S_n(\omega)>0\}$. Как и в симметричном случае (см. § 11), имеем при $0<n_1<n_2<n$, $m_1, m_2>0$

$$p_{Y_{n(2)}|Y_{n(1)}}(m_2|m_1)=\mathbf{P}(\omega : S_{n_2-n_1}(\omega)=m_2-m_1)\cdot\mathbf{P}(\omega : S_{n-n_2}(\omega)>-m_2)\times$$
$$\times\mathbf{P}(\omega : S_{n-n_1}(\omega)>-m_1)^{-1},\quad n(i)\equiv n_i.$$

С учетом того, что сл. в. $(S_n+n)/2$ имеет биномиальное распределение $p_n(k;\,p)$, выводим для переходных вероятностей нужную асимптотику, полагая $n_i=[nt_i]$, $m_i=[x_i\sqrt{n}]$, $i=1, 2$, $0<t_1<t_2<1$, и используя 12.4:

$$C_{n_2-n_1}^{(n_2-n_1+m_2-m_1)/2}p^{(n_2-n_1+m_2-m_1)/2}(1-p)^{(n_2-n_1-m_2+m_1)/2}\times$$
$$\times C_{n-n_2}^{(n-n_2-m_2)/2}p^{(n-n_2-m_2)/2}(1-p)^{(n-n_2+m_2)/2}\cdot\frac{2^{-1}(1-p)}{2^{-1}-p}\times$$
$$\times\left(C_{n-n_1}^{(n-n_1-m_1)/2}p^{(n-n_1-m_1)/2}(1-p)^{(n-n_1+m_1)/2}\cdot\frac{2^{-1}(1-p)}{2^{-1}-p}\right)^{-1}=$$
$$=C_{n_2-n_1}^{(n_2-n_1+m_2-m_1)/2}C_{n-n_2}^{(n-n_2-m_2)/2}\left(C_{n-n_1}^{(n-n_1-m_1)/2}\right)^{-1},$$

что совпадает с выражением из 11.18 для условного симметричного случайного блуждания с условием $S_n=0$.]

§ 13.

СХОДИМОСТЬ РАСПРЕДЕЛЕНИЙ, ПРЕОБРАЗОВАНИЕ ЛАПЛАСА И ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ

13.1. Задача. Рассмотрим последовательности ф. р.: (I) $F_n(x)=\Phi(x/\sigma_n)$, II) $F_n(x)=1-e^{-x/\sigma_n}$, $x>0$, $F_n(x)=0$, $x\leqslant 0$, где $\sigma_n>0$, $\sigma_n\to 0$ при $n\to\infty$, $\Phi(x)$ — стандартная нормальная ф. р. Вычислить $\lim\limits_{n\to\infty} F_n(x)$. Какое распределение вероятностей следует считать предельным для последовательности F_n?

[(I) $\lim\limits_{n\to\infty} F_n(x)=E(x)$, $E(x)=1$ при $x>0$, $E(x)=0$ при $x<0$, $E(0)=1/2$; (II) $\lim\limits_{n\to\infty} F_n(x)=H(x)$, $H(x)=1$ при $x>0$, $H(x)=0$ при $x\leqslant 0$. Последовательность вероятностных мер $\mathbf{P}_n$ с ф. р. F_n в обоих случаях такова, что $\mathbf{P}_n(I_{a,b})\to 1$ при $a<0<b$, $\mathbf{P}_n(I_{a,b})\to 0$ при $a<b<0$ или $0<a<b$, т. е. вероятностные массы, распределенные по закону $\mathbf{P}_n$, стягиваются в точку нуль. Хотя в обоих случаях $F_n(0)\not\to F(0)$, где $F(x)$ — функция распределения меры $\mathbf{P}$, целиком сосредоточенной в нуле, и, следовательно, $\mathbf{P}_n(I_{a,b})\not\to\mathbf{P}(I_{a,b})$, если одно из чисел a или b совпадает с нулем ($a<b$), все же представляется разумным считать, что мера $\mathbf{P}$ является пределом последовательности мер $\mathbf{P}_n$.]

13.2. Задача. Пусть $x_1, \ldots, x_m$ — произвольный набор попарно различных чисел, $p_1, \ldots, p_m$ — распределение вероятностей на множестве $\{x_1, \ldots, x_m\}$. Составим *смесь* ф. р. $F_n(x-x_k)$, $k=1, \ldots, m$, где ф. р. F_n одного из двух типов, рассмотренных в 13.1:

$$G_n(x)=p_1F_n(x-x_1)+\ldots+p^mF_n(x-x_m).$$

Вычислить $\lim\limits_{n\to\infty} G_n(x)$. Какое распределение вероятностей следует считать предельным?

[Для любого достаточно малого $\varepsilon>0$ $\mathbf{P}_n((x_k-\varepsilon, x_k+\varepsilon))\to p_k$, $k=1, \ldots, m$, $n\to\infty$, и в качестве предельного следует принять дискретное распределение $p(x_k)=p_k$, $k=1, \ldots, m$, с ф. р.

$$G(x)=\sum_{k=1}^{m} p_k I_{(-\infty, x_k]}(x).$$

Отметим, что в точках разрыва предельной ф. р. $G(x)$ соотношение $\lim\limits_{n\to\infty} G_n(x)=G(x)$ не выполняется.]

Скажем, что последовательность распределений вероятностей $\mathbf{P}_n$ на $(R, \mathscr{B})$ и последовательность их ф. р. F_n *слабо сходятся* к распределению $\mathbf{P}$ на $(R, \mathscr{B})$ и соответственно к ф. р. $F(x)=\mathbf{P}((-\infty, x])$, если соотношение

$$\lim_{n\to\infty} F_n(x)=F(x)$$

выполняется во всех точках непрерывности ф. р. $F(x)$. Для слабой сходимости употребляется обозначение $\mathbf{P}_n \Rightarrow \mathbf{P}$, $F_n \Rightarrow F$. Если $F_{X_n} \Rightarrow F_X$, то говорят, что последовательность сл. в. X_n, $n=1, 2, \ldots$, сходится по распределению к сл. в. X, и пишут $X_n \xrightarrow{\mathscr{D}} X$ (distribution — распределение). Подчеркнем, что сходимость случайных величин по распределению — это сходимость их распределений, и она не имеет непосредственного отношения к сходимости случайных величин как функций на пространстве элементарных событий, более того, сл. в. X, X_1, X_2, ... могут быть определены каждая на своем вероятностном пространстве .

Пусть сл. в. $X(\omega)$, $X_n(\omega)$, $n=1, 2, \ldots$, определены на едином вероятностном пространстве $(\Omega, \mathscr{B}, \mathbf{P})$. Говорят, что последовательность X_n, $n=1, 2, \ldots$, *сходится к X с вероятностью* 1 или *почти наверное* (п. н.), если

$$\mathbf{P}(\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)) = 1,$$

и X_n *сходится* к X *по вероятности*, если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) = 0.$$

В первом случае употребляют обозначение $\lim\limits_{n\to\infty} X_n(\omega) = X(\omega)$ п. н. (или $X_n(\omega) \to X(\omega)$ при $n \to \infty$ п. н.), а во втором — $X_n \xrightarrow{\mathbf{P}} X$. В 12.14 показано, что из сходимости п. н. вытекает сходимость по вероятности.

13.3. Задача. Допустим, что сл. в. $X_n(\omega)$, $n=1, 2, \ldots$, определены на едином вероятностном пространстве. Показать, что $X_n \xrightarrow{\mathscr{D}} a = \mathrm{const}$ тогда и только тогда, когда $X_n \xrightarrow{P} a$, $n \to \infty$.
[Пусть $X_n \xrightarrow{\mathscr{D}} a$, тогда $F_n(x) \to 0$ при $x < a$ и $F_n(x) \to 1$ при $x > a$. Отсюда заключаем, что

$$\mathbf{P}(\omega : a - \varepsilon < X_n(\omega) \leqslant a + \varepsilon) = F_n(a+\varepsilon) - F_n(a-\varepsilon) \to 1, \quad n \to \infty,$$

т. е. $X_n \xrightarrow{P} X$. Обратно, из $X_n \xrightarrow{P} a$ вытекает, что $F_n(a+\varepsilon) \to 1$, $F_n(a-\varepsilon) \to 0$ при любом $\varepsilon > 0$, $n \to \infty$, что и означает сходимость по распределению $X_n \xrightarrow{\mathscr{D}} a$.]

Как видно из 13.3, сходимость по вероятности к константе можно определить для последовательности случайных величин, определенных на разных вероятностных пространствах, и притом она остается эквивалентной сходимости по распределению. Это допущение, позволяющее сходимость по распределению рассматривать на языке случайных величин, бывает часто полезным.

13.4. Задача. Показать, что из сходимости по вероятности вытекает сходимость по распределению.
[Предположим, что $X_n \xrightarrow{P} X$ и x — точка непрерывности ф. р. F. Для любого $\varepsilon > 0$ имеем

$$F_n(x)=\mathbf{P}(\omega:X_n(\omega)\leqslant x)=\mathbf{P}(\omega:X_n(\omega)\leqslant x,\ |X_n(\omega)-X(\omega)|\leqslant\varepsilon)+$$
$$+\mathbf{P}(\omega:X_n(\omega)\leqslant x,\ |X_n(\omega)-X(\omega)|>\varepsilon).$$

Если $X_n\leqslant x$, $|X_n-X|\leqslant\varepsilon$, то $X\leqslant x+\varepsilon$ и потому

$$F_n(x)\leqslant\mathbf{P}(\omega:X(\omega)\leqslant x+\varepsilon)+\mathbf{P}(\omega:|X_n(\omega)-X(\omega)|>\varepsilon).$$

Аналогично

$$F_n(x)\geqslant\mathbf{P}(\omega:X_n(\omega)\leqslant x,\ |X_n(\omega)-X(\omega)|\leqslant\varepsilon)\geqslant$$
$$\geqslant\mathbf{P}(\omega:X(\omega)+\varepsilon\leqslant x,\ |X_n(\omega)-X(\omega)|\leqslant\varepsilon)\geqslant$$
$$\geqslant\mathbf{P}(\omega:X(\omega)\leqslant x-\varepsilon)-\mathbf{P}(\omega:|X_n(\omega)-X(\omega)|>\varepsilon).$$

Переходя в полученных неравенствах к пределу по $n\to\infty$:

$$F(x-\varepsilon)\leqslant\varliminf_{n\to\infty}F_n(x)\leqslant\varlimsup_{n\to\infty}F_n(x)\leqslant F(x+\varepsilon),$$

а затем по $\varepsilon\to0$, получаем требуемый результат.]

13.5. Задача. Пусть сл. в. X_n, Y_n определены на едином вероятностном пространстве при каждом $n=1, 2, \ldots$, $X_n\xrightarrow{\mathscr{D}}X$, $Y_n\xrightarrow{\mathscr{D}}0$. Показать, что $X_n+Y_n\xrightarrow{\mathscr{D}}X$.

$$[\mathbf{P}(\omega:X_n(\omega)+Y_n(\omega)\leqslant x)=\mathbf{P}(\omega:X_n(\omega)+Y_n(\omega)\leqslant x,\ |Y_n(\omega)|\leqslant\varepsilon)+$$
$$+\mathbf{P}(\omega:X_n(\omega)+Y_n(\omega)\leqslant x,\ |Y_n(\omega)|>\varepsilon),$$
$$F_{X_n+Y_n}(x)\leqslant F_{X_n}(x+\varepsilon)+\mathbf{P}(\omega:|Y_n(\omega)|>\varepsilon),$$
$$F_{X_n+Y_n}(x)\geqslant F_{X_n}(x-\varepsilon)-\mathbf{P}(\omega:|Y_n(\omega|>\varepsilon).$$

Пусть x — точка непрерывности ф. р. $F(x)$. Учитывая, что любая ф. р. имеет не более чем счетное число точек разрыва (см. 6.23), выберем последовательность $\varepsilon_k>0$, $\varepsilon_k\to0$ при $k\to\infty$ таким образом, чтобы ф. р. F была непрерывна во всех точках вида $x\pm\varepsilon_k$, $k=1, 2, \ldots$. Переходя в полученных неравенствах к пределу по $n\to\infty$:

$$F_X(x-\varepsilon_k)\leqslant\varliminf_{n\to\infty}F_{X_n+Y_n}(x)\leqslant\varlimsup_{n\to\infty}F_{X_n+Y_n}(x)\leqslant F(x+\varepsilon_k),$$

а затем по $k\to\infty$, получаем требуемое утверждение.]

13.6. Задача. Последовательность сл. в. X_n, $n=1, 2, \ldots$, называется *ограниченной по вероятности*, если для любого $\varepsilon>0$ найдется такое $A>0$, что $\mathbf{P}(\omega:|X_n(\omega)|>A)<\varepsilon$ для всех достаточно больших n. Показать, что если $X_n\xrightarrow{\mathscr{D}}X$, то последовательность X_n ограничена по вероятности.

$[\mathbf{P}_n(\omega:-A<X_n(\omega)\leqslant A)=F_n(A)-F_n(-A)\to F(A)-F(-A)$,

если только $F(x)$ непрерывна в точках $\pm A$. По заданному $\varepsilon>0$ выберем $A>0$ настолько большим, что

$$1-F(A)+F(-A)<\varepsilon.$$

Увеличивая A, добьемся того, что точки $\pm A$ будут точками непрерывности ф. р. F (ср. 13.5). Дальнейшее очевидно.]

13.7. Задача. Пусть сл. в. X_n, Y_n определены на едином вероятностном пространстве при каждом $n=1, 2, \ldots$, $X_n \xrightarrow{\mathscr{L}} 0$, Y_n — ограничена по вероятности (см. 13.6). Показать, что $X_nY_n \xrightarrow{\mathscr{L}} 0$.

$$[\mathbf{P}(\omega: |X_n(\omega)Y_n(\omega)|>\varepsilon) \leqslant \mathbf{P}(\omega: |X_n(\omega)Y_n(\omega)|>\varepsilon,\ |Y_n(\omega)|\leqslant A)+$$
$$+\mathbf{P}(\omega: |X_n(\omega)Y_n(\omega)|>\varepsilon,\ |Y_n(\omega)|>A) \leqslant \mathbf{P}(\omega: |X_n(\omega)|>\varepsilon/A)+$$
$$+\mathbf{P}(\omega: |Y_n(\omega)|>A) \leqslant \mathbf{P}(\omega: |X_n(\omega)|>\varepsilon/A)+\varepsilon,$$

где последнее неравенство справедливо в предположении, что A, зависящее от ε, выбрано достаточно большим. В таком случае

$$\varlimsup_{n\to\infty} \mathbf{P}(\omega: |X_n(\omega)Y_n(\omega)|>\varepsilon) \leqslant \varepsilon,$$

что и доказывает утверждение.]

13.8. Задача. Пусть сл. в. X_n, Y_n определены на едином вероятностном пространстве при каждом $n=1, 2, \ldots$, $X_n \xrightarrow{\mathscr{L}} X$, $Y_n \xrightarrow{\mathscr{L}} 1$. Показать, что $X_nY_n \xrightarrow{\mathscr{L}} X$.
[$X_nY_n=X_n(Y_n-1)+X_n$, и остается воспользоваться результатами 13.5—13.7.]

13.9. Задача. Пусть функция $g(x, y, z)$ непрерывна в точке (a, b, c), $X_n \xrightarrow{P} a$, $Y_n \xrightarrow{P} b$, $Z_n \xrightarrow{P} c$, где $X_n(\omega)$, $Y_n(\omega)$, $Z_n(\omega)$ определены на едином вероятностном пространстве. Показать, что $g(X_n, Y_n, Z_n) \xrightarrow{P} g(a, b, c)$.
[Для любого $\varepsilon>0$ найдется $\delta>0$, такое, что

$$\{x, y, z: |x-a|<\delta,\ |y-b|<\delta,\ |z-c|<\delta\} \subseteq$$
$$\subseteq \{x, y, z: |g(x, y, z)-g(a, b, c)|<\varepsilon\},$$

откуда получаем

$$\{\omega: |g(X_n(\omega), Y_n(\omega), Z_n(\omega))-g(a, b, c)| \geqslant \varepsilon\} \subseteq$$
$$\subseteq \{\omega: |X_n(\omega)-a| \geqslant \delta\} \cup \{\omega: |Y_n(\omega)-b| \geqslant \delta\} \cup \{\omega: |Z_n(\omega)-c| \geqslant \delta\}.$$

Переход к вероятностям заканчивает доказательство.]

13.10. Задача. Рассмотрим последовательность распределений вероятностей: (I) пуассоновских с параметром λ_n, $n=1, 2, \ldots$; (II) гамма — с плотностями $f_n(x)=(x^{\lambda_n}/\Gamma(\lambda_n+1))e^{-x}$, $x>0$. Вывести из 12.5 и 12.7, что если $\lambda_n\to\infty$ при $n\to\infty$, то последовательность ф. р. $F_n^*(x)=F_n(\lambda_n+x\sqrt{\lambda_n})$ в обоих случаях слабо сходится к стандартному нормальному закону.

(I) Пусть сл. в. X_n и Y_n независимы и имеют пуассоновские распределения с параметрами $[\lambda_n]$ и $\lambda_n-[\lambda_n]$ соответственно (полагаем пуассоновскую сл. в. с параметром 0 равной нулю с вероятностью 1). Тогда X_n+Y_n имеет пуассоновское распределение с параметром λ_n (см. 4.7). В 12.5 показано, что $(X_n-[\lambda_n])/[\lambda_n]^{1/2}$ сходится по распределению к стандартной нормальной величине. Тогда

$$\frac{X_n+Y_n-\lambda_n}{(\lambda_n)^{1/2}}=\frac{X_n-[\lambda_n]}{[\lambda_n]^{1/2}}\cdot\frac{[\lambda_n]^{1/2}}{\lambda_n^{1/2}}+\frac{Y_n}{\lambda_n^{1/2}}+\frac{[\lambda_n]-\lambda_n}{\lambda_n^{1/2}},$$

и остается воспользоваться результатами 13.5, 13.8, заметив, что сходящиеся числовые последовательности могут рассматриваться как сходящиеся по распределению постоянные случайные величины. Аналогично устанавливается утверждение (II) со ссылкой на 4.19 вместо 4.7.]

13.11. Пример (асимптотическая нормальность выборочных квантилей). Пусть $Y_1, \ldots, Y_n$ — независимые равномерно распределенные на $[0, 1]$ случайные величины, $Y_{k,n}$ — k-я порядковая статистика, $Y_{[np],n}$, $0<p<1$, — выборочная квантиль. В 12.8 было установлено, что $(n/(p(1-p)))^{1/2}(Y_{[np],n}-p)$ сходится по распределению к стандартной нормальной величине:

$$Z_n=\sqrt{n/p(1-p)}\,(Y_{[np],n}-p)\xrightarrow{\mathscr{L}}Z,\quad n\to\infty,$$

где Z — стандартная нормальная величина. Отметим, что отсюда следует (см. 13.7, 13.6):

$$Y_{[np],n}-p=Z_n\cdot\sqrt{p(1-p)/n}\xrightarrow{P}0,\quad n\to\infty.$$

Последовательность независимых сл. в. $X_1, \ldots, X_n$ с произвольной непрерывной ф. р. $F(x)$ можно образовать, полагая (см. 12.10)

$$X_i=G(Y_i),\quad G(y)\equiv F^{-1}(y)=\sup\{x: F(x)=y\}.$$

Выведем отсюда предельную теорему для выборочных квантилей $X_{[np],n}=G(Y_{[np],n})$, предполагая, что функция $G(y)$ имеет производную в точке p, т. е. $F(x)$ имеет положительную производную в точке $x_p=G(p)$ — p-квантили ф. р. $F(x)$. По определению производной

$$G(y)-G(p)=(y-p)\,(1/f(x_p)+\alpha(y, p)),$$

где $f(x)=(d/dx)F(x)$, $\alpha(y, p)\to0$ при $y\to p$. Совершая подстановку $y=Y_{[np],n}$, запишем

$$X_{[np],n}-x_p=\left(Z_n/\sqrt{n/p(1-p)}\right)(1/f(x_p)+\alpha(Y_{[np],n}, p))$$

или

$$\left(f(x_p)/\sqrt{p(1-p)}\right)\sqrt{n}\,(X_{[np],n}-x_p)=Z_n\,(1+f(x_p)\,\alpha(Y_{[np],n}, p)).$$

Доопределив функцию $\alpha(y, p)$ в точке $y=p$ по непрерывности, учитывая, что $Y_{[np],n} \xrightarrow{P} 0$ и применяя 13.8, заключаем, что

$$\left(f(x_p)/\sqrt{p(1-p)}\right)\sqrt{n}\,(X_{[np],n}-x_p) \xrightarrow{\mathcal{D}} Z, \quad n\to\infty.$$

Приведем другое доказательство предельной теоремы для последовательности $Y_{[np,]n}$, $n=1, 2, \ldots$, выборочных квантилей равномерного распределения на $[0, 1]$. Пусть сл. в. $V_1, \ldots, V_{n+1}$ — независимые, стандартные экспоненциальные, $\tau_k=V_1+\ldots+V_k$, $k\leqslant n+1$. Тогда (см. 4.25) сл. в. $\tau_1/\tau_{n+1}, \ldots, \tau_n/\tau_{n+1}$ имеют такое же совместное распределение, как порядковые статистики $Y_{1,n}, \ldots$ $\ldots, Y_{n,n}$ из равномерного распределения. Отсюда получаем, что распределение $X_{k,n}$ такое же, как у τ_k/τ_{n+1}. Проводя элементарные преобразования, имеем

$$\sqrt{n}\left(\frac{\tau_k}{\tau_{n+1}}-p\right)=\sqrt{n}\,\frac{(1-p)\tau_k-p(\tau_{n+1}-\tau_k)}{\tau_{n+1}}=\sqrt{n}\,\frac{k-(n+1)p}{\tau_{n+1}}+$$

$$+\frac{n}{\tau_{n+1}}\left((1-p)\frac{\tau_k-k}{\sqrt{k}}\frac{\sqrt{k}}{\sqrt{n}}-p\,\frac{(\tau_{n+1}-\tau_k)-(n+1-k)}{\sqrt{n+1-k}}\cdot\frac{\sqrt{n+1-k}}{\sqrt{n}}\right).$$

Сл. в. $\tau_k=V_1+\ldots+V_k$ и $\tau_{n+1}-\tau_k=V_{k+1}+\ldots+V_{n+1}$ независимы и, как установлено в 12.7, последовательность сл. в. $\overset{\circ}{\tau}_m=(\tau_m-m)/\sqrt{m}$, $m\to\infty$, имеет стандартное нормальное предельное распределение. Положим

$$U_n=(1-p)(\tau_k-k)/\sqrt{n}, \quad W_n=-p(\tau_{n+1}-\tau_k-n-1+k)/\sqrt{n}.$$

С учетом сказанного, принимая во внимание 13.8, выводим из 12.7, что при $n\to\infty$, $k=np$

$$f_{U_n,W_n}(u, w)\to\varphi\left(u/\left(\sqrt{p}\,(1-p)\right)\right)\varphi\left(w/\left(p\sqrt{1-p}\right)\right)$$

равномерно по (u, w) в любой ограниченной области. Легко видеть, что тогда сл. в. U_n+W_n имеют в пределе нормальное распределение со средним нуль и дисперсией $p(1-p)$. Наконец,

$$\sqrt{n}\,(\tau_k/\tau_{n+1}-p)=(n/\tau_{n+1})\,(U_n+W_n+(k-(n+1)p)/\sqrt{n}),$$

и остается заметить, что $n/\tau_{n+1}\xrightarrow{P}1$.

13.12. Задача. Пусть X, X_n, $n=1, 2, \ldots$, — сл. в. с целыми неотрицательными значениями, $X_n\xrightarrow{\mathcal{D}}X$. Показать, что сходятся производящие функции

$$Mt^{X_n}=\mathcal{P}_{X_n}(t)\to\mathcal{P}_X(t)=Mt^X, \quad 0\leqslant t\leqslant 1.$$

[Легко видеть, что сходимость по распределению $X_n\xrightarrow{\mathcal{D}}X$ имеет место тогда и только тогда, когда $p_{X_n}(k)\to p_X(k)$, $n\to\infty$, при всех $k=0, 1, 2, \ldots$. Так как $|p_{X_n}(k)-p_X(k)|\leqslant 2$, то при $0\leqslant t\leqslant 1$ получаем

$$|\mathcal{P}_{X_n}(t)-\mathcal{P}_X(t)|\leqslant\sum_{k=0}^{r-1}|p_{X_n}(k)-p_X(k)|+2t^r/(1-t).$$

Устремляя к бесконечности n, а затем r, получаем требуемое утверждение.]

13.13.* Задача. Установить утверждение, обратное к 13.12.

Покажем сначала, что если произвольная последовательность вероятностных п. ф. $\mathcal{P}^{(n)}(t)$ сходится при $0<t\leqslant 1$, то она сходится и при $t=0$. Запишем неравенство

$$p_0^{(n)}\leqslant\mathcal{P}^{(n)}(t)=p_0^{(n)}+p_1^{(n)}t+p_2^{(n)}t^2+\ldots\leqslant p_0^{(n)}+t/(1-t),\ 0<t<1.$$

Устремляя n к ∞:

$$\mathcal{P}(t)-t/(1-t)\leqslant\varliminf_{n\to\infty}p_0^{(n)}\leqslant\varlimsup_{n\to\infty}p_0^{(n)}\leqslant\mathcal{P}(t)$$

а затем t к 0, получаем требуемое.

Полагая $t=0$, из соотношения $\mathcal{P}_{X_n}(t)\to\mathcal{P}_X(t)$, $n\to\infty$, выводим, что $p_{X_n}(0)\to p_X(0)$. Образуем вероятностные п. ф.

$$\mathcal{P}^{(n)}(t)=\frac{\mathcal{P}_{X_n}(t)-p_{X_n}(0)}{t(1-p_{X_n}(0))},\quad \mathcal{P}(t)=\frac{\mathcal{P}_X(t)-p_X(0)}{t(1-p_X(0))},$$

доопределяя их при $t=0$ по непрерывности. Очевидно, что $\mathcal{P}^{(n)}(t)\to\mathcal{P}(t)$ при $0<t\leqslant 1$, а следовательно, $\mathcal{P}^{(n)}(0)\to\mathcal{P}(0)$, $n\to\infty$. Но $\mathcal{P}^{(n)}(0)=p_{X_n}(1)/(1-p_{X_n}(0))$, $\mathcal{P}(0)=p_X(1)/(1-p_X(0))$, и потому $p_{X_n}(1)\to p_X(1)$, $n\to\infty$. По индукции устанавливаем сходимость $p_{X_n}(k)\to p_X(k)$, $n\to\infty$, при всех k.]

13.14. Задача. Применить 13.13 к доказательству теоремы Пуассона: $p_n(k;\ p)\to(\lambda^k/k!)e^{-\lambda}$ при $n\to\infty$, $np\to\lambda>0$.

[Используя п. ф. биномиального и пуассоновского распределений из 4.10, имеем

$$\mathcal{P}_{n,p}(t)=(1-p(1-t))^n\to e^{-\lambda(1-t)}=\mathcal{P}(t).$$

Остается применить 13.13.]

Преобразованием Лапласа неотрицательной сл. в. X и ее распределения вероятностей называется функция

$$\psi_X(\lambda)=\mathrm{M}e^{-\lambda X},\quad \lambda\geqslant 0.$$

Функция $\psi_X(\lambda)$ определена при всех $\lambda\geqslant 0$, так как математическое ожидание ограниченной случайной величины существует; при этом $0<\psi_X(\lambda)\leqslant 1$ (см. 7.23). Если сл. в. X принимает лишь целые (неотрицательные) значения, то

$$\psi_X(\lambda)=\mathcal{P}_X(e^{-\lambda}),$$

где $\mathcal{P}_X(t)$ — производящая функция X. Если сл. в. X дискретна, то

$$\psi_X(\lambda)=\sum_{\{x\}} e^{-\lambda x} p_X(x),$$

где суммирование ведется по множеству всех значений сл. в. X. Если сл. в. X имеет плотность $f_X(x)$, то рассуждения задачи 7.29, примененные к функции $e^{-\lambda x}$ вместо функции x^r, дают следующую формулу:

$$\psi_X(\lambda)=\int_0^{\infty} e^{-\lambda x} f_X(x)\,dx.$$

13.15. Задача. Вычислить преобразование Лапласа для гамма-распределения с плотностью $f(x)=(1/\Gamma(\alpha))\,x^{\alpha-1}e^{-x}$, $x>0$.

$$\left[\ \psi(\lambda)=\Gamma(\alpha)^{-1}\int_0^{\infty} x^{\alpha-1}e^{-(\lambda+1)x}dx=(\lambda+1)^{-\alpha},\ \alpha>0.\ \right]$$

В общем случае преобразование Лапласа $\psi_X(\lambda)$ можно представить *интегралами Римана — Стильтьеса* или *Лебега — Стильтьеса*

$$\psi_X(\lambda)=\int_0^{\infty} e^{-\lambda x}\,dF_X(x).$$

Для любой неубывающей функции $F(x)$, $-\infty<x<\infty$, теория интегралов Римана — Стильтьеса

$$\int_a^b g(x)\,dF(x)$$

почти дословно повторяет теорию интеграла Римана (который соответствует частному случаю $F(x)=x$). Именно, рассматривают интегральные суммы

$$\sum_{i=1}^{n} g(t_i)(F(x_i)-F(x_{i-1})),\quad t_i\in[x_{i-1},\ x_i],\quad a=x_0<x_1<\ldots<x_n=b$$

и их пределы при диаметре разбиения $d=\max\limits_{(i)}(x_i-x_{i-1})$, стремящемся к нулю, либо суммы Дарбу

$$\sum_{i=1}^{n} m_i(F(x_i)-F(x_{i-1})),\quad \sum_{i=1}^{n} M_i(F(x_i)-F(x_{i-1})),$$

$$m_i=\inf_{\{x\in[x_{i-1},x_i]\}} g(x),\quad M_i=\sup_{\{x\in[x_{i-1},x_i]\}} g(x).$$

Для интегрируемой по Риману — Стильтьесу функции точная верхняя грань всех нижних сумм совпадает с точной нижней гранью всех верхних и совпадает с пределом любой последовательности интегральных сумм с диаметром разбиений, стремящимся к нулю. Все элементарные свойства интеграла Римана выполняются и для интеграла Римана — Стильтьеса. К классу интегрируемых функций принадлежат, в частности, кусочно-непрерывные функции, кусочно-монотонные, какова бы ни была функция $F(x)$ (неубывающая). Интеграл по бесконечному промежутку определяется как предел интегралов по расширяющимся конечным промежуткам.

Интегралы Лебега — Стильтьеса могут быть введены по схеме построения математического ожидания из § 7, если числовую прямую R принять за пространство элементарных событий, совокупность $\mathscr{B}$ всех борелевских подмножеств — за класс событий, меру $\mathbf{P}_F$ задать с помощью функции распределения $F_X(x)$. Тогда для любой действительной функции $g(x)$ на R, такой, что $\{x : g(x) \in \in I_{a,b}\} \in \mathscr{B}$ для любого промежутка $I_{a,b}$, интеграл Лебега — Стильтьеса

$$\int g(x)\, dF_X(x)$$

определяется как математическое ожидание от случайной величины $g(x)$ на вероятностном пространстве $(R, \mathscr{B}, \mathbf{P}_F)$. Более обстоятельно теория интеграла Лебега — Стильтьеса изложена в § 10. Здесь ограничимся некоторыми замечаниями, используя частный вид функции $g(x) = e^{-\lambda x}$.

Положим $X^{(n)}(\omega) = kn^{-1}$ при $X(\omega) \in (kn^{-1}, (k+1)n^{-1}]$, $k = 0, 1, 2, \ldots$, и $X^{(n)}(\omega) = 0$ при $X(\omega) = 0$. В соответствии с 7.19 имеем

$$\mathbf{M}e^{-\lambda X} = \lim_{n\to\infty} \mathbf{M}e^{-\lambda X^{(n)}},$$

$$\mathbf{M}e^{-\lambda X^{(n)}} = 1 \cdot (F_X(n^{-1}) - F_X(0-)) + \\ + \sum_{k=1}^{\infty} e^{-\lambda k n^{-1}} (F_X((k+1)n^{-1}) - F_X(kn^{-1})).$$

Последнее выражение представляет собой интегральную сумму интеграла Римана — Стильтьеса по бесконечному промежутку $[0, +\infty)$. Так как интегральная сумма по промежутку $[A, \infty)$ не превосходит, очевидно, $1 - F_X(A-)$, что может быть сделано меньше любого $\varepsilon > 0$ за счет выбора A достаточно большим, то отсюда следует, что преобразование Лапласа представимо в виде интеграла Римана — Стильтьеса.

Отметим следующие простые свойства преобразования Лапласа:

$$\psi_{aX+b}(\lambda) = e^{-\lambda b}\psi_X(\lambda a), \quad \psi_{X_1+X_2}(\lambda) = \psi_{X_1}(\lambda) \cdot \psi_{X_2}(\lambda)$$

для независимых X_1, X_2, которые непосредственно вытекают из определения $\psi_X(\lambda)$ и свойств математического ожидания. Дальнейшие свойства устанавливаются ниже.

13.16. Задача. Показать, что k-я производная от $\psi_X(\lambda)$ существует и выражается формулой

$$\psi_X^{(k)}(\lambda)=(-1)^k\int_0^\infty y^k e^{-\lambda y}\,dF_X(y),\quad \lambda>0.$$

[Пусть для простоты $k=1$. Поскольку

$$(e^{-(\lambda+\Delta\lambda)y}-e^{-\lambda y})/\Delta\lambda\to -ye^{-\lambda y},\quad \Delta\lambda\to 0,$$

равномерно по $y\geqslant 0$, то результат вытекает из легко устанавливаемого свойства интеграла Римана — Стильтьеса: возможности перейти к пределу под знаком интеграла в случае равномерной сходимости. Другой способ — воспользоваться теорией интеграла Лебега из § 10.]

13.17. Задача. Установить, используя неравенство Чебышева, что

$$\sum_{\{k\leqslant\lambda(t+\delta)\}}\frac{(\lambda t)^k}{k!}e^{-\lambda t}\to\begin{cases}1 & \text{при } \delta>0,\\ 0 & \text{при } \delta<0,\end{cases}$$

при $\lambda\to\infty$ равномерно по $t\geqslant 0$, δ, таким, что $t\delta^{-2}$ ограничено. [Функция

$$F_{\lambda,t}(x)=\sum_{\{k\leqslant x\}}((\lambda t)^k/k!)\,e^{-\lambda t}$$

— ф. р. пуассоновской сл. в. X с параметром λt. Результат легко следует из неравенства Чебышева

$$\mathbf{P}(\omega:\ |X(\omega)-\lambda t|>\lambda\delta)\leqslant\lambda t/(\lambda\delta)^2=\lambda^{-1}t\delta^{-2},\quad \delta>0,$$

так что

$$F_{\lambda,t}(\lambda t-\lambda\delta)\leqslant\lambda^{-1}t\delta^{-2},\quad F_{\lambda,t}(\lambda t+\lambda\delta)\geqslant 1-\lambda^{-1}t\delta^{-2}.]$$

Перепишем соотношение 13.17 в виде

$$h_{\lambda,x}(y)\equiv\sum_{\{k\leqslant\lambda x\}}\frac{(\lambda y)^k}{k!}e^{-\lambda y}\to\begin{cases}1 & \text{при } y<x,\\ 0 & \text{при } y>x,\end{cases}$$

при $\lambda\to\infty$ равномерно по x, $y\geqslant 0$, таким, что $y(x-y)^{-2}$ ограничено. Отметим, что $0\leqslant h_{\lambda,x}(y)\leqslant 1$. Воспользуемся полученным соотношением для доказательства того, что преобразование Лапласа однозначно определяет распределение вероятностей. Домножив соотношения 13.16 на $(-\lambda)^k/k!$ и просуммировав по $k\leqslant\lambda x$, получим.

$$\sum_{\{k\leqslant\lambda x\}}\frac{(-\lambda)^k}{k!}\psi_X^{(k)}(\lambda)=\int_0^\infty h_{\lambda,x}(y)\,dF_X(y).$$

Покажем, что при $\lambda \to \infty$ интеграл справа стремится к $F_X(x)$, если только x есть точка непрерывности ф. р. F_X.

Как было установлено, функция $h_{\lambda,x}(y) \to 1$ равномерно на любом отрезке $[0, x-\varepsilon]$, $\varepsilon > 0$, и $h_{\lambda,x} \to 0$ равномерно на $[x+\varepsilon, \infty)$. Учитывая, что (см. определение интеграла Римана — Стильтьеса)

$$\int_{[a,b]} 1 \cdot dF_X(y) = F_X(b) - F_X(a-), \quad \int_{(a,\infty)} 1 \cdot dF_X(y) = 1 - F_X(a),$$

получаем, что

$$\int_{[0,x-\varepsilon]} h_{\lambda,x}(y)\,dF_X(y) \to F_X(x-\varepsilon), \quad \lambda \to \infty,$$

$$\int_{(x+\varepsilon,\infty)} h_{\lambda,x}(y)\,dF_X(y) \to 0, \quad \lambda \to \infty,$$

$$\int_{(x-\varepsilon,x+\varepsilon)} h_{\lambda,x}(y)\,dF_X(y) \leqslant F_X(x+\varepsilon) - F_X(x-\varepsilon).$$

Устремляя ε к 0, получаем следующую *формулу обращения* для преобразования Лапласа:

$$F_X(x) = \lim_{\lambda \to \infty} \sum_{\{k \leqslant \lambda x\}} \frac{(-\lambda)^k}{k!} \psi_X^{(k)}(\lambda)$$

для любой точки непрерывности x функции F_X. Поскольку функция распределения определяется своими значениями в точках непрерывности, а распределение вероятностей определяется своей функцией распределения, то утверждение установлено.

13.18. Задача. Пусть последовательность ф. р. $F_n(x)$ слабо сходится к ф. р. $F(x)$, $h(x)$ — произвольная ступенчатая функция на отрезке $[a, b]$:

$$h(x) = \sum_{i=1}^{n} c_i I_{(x_{i-1}, x_i]}(x),$$

$$a = x_0 < x_1 < \ldots < x_n = b.$$

Предположим, что скачки функции $h(x)$ приходятся на точки непрерывности ф. р. $F(x)$. Тогда имеет место соотношение

$$\lim_{n \to \infty} \int h(x)\,dF_n(x) = \int h(x)\,dF(x).$$

$$\Big[\int h(x)\,dF_n(x) = \sum_{i=1}^{n} c_i \int I_{(x_{i-1}, x_i]}(x)\,dF_n(x) =$$

$$= \sum_{i=1}^{n} c_i (F_n(x_i) - F_n(x_{i-1})) \to \sum_{i=1}^{n} c_i (F(x_i) - F(x_{i-1})) = \int h(x)\,dF(x). \Big]$$

13.19. Задача. Пусть $F_n \Rightarrow F$ и функция $g(x)$ непрерывна на $[a, b]$. Показать, что

$$\int_a^b g(x)\,dF_n(x) \to \int_a^b g(x)\,dF(x), \quad n \to \infty.$$

[Для любого $\varepsilon > 0$ выберем ступенчатую функцию $h_\varepsilon(x)$ на $[a, b]$, такую, что $|g(x) - h_\varepsilon(x)| < \varepsilon$ при $x \in [a, b]$, а ее скачки приходятся на точки непрерывности функции $F(x)$. В таком случае

$$\overline{\lim_{n\to\infty}} \int_a^b g(x)\,dF_n(x) \leqslant \overline{\lim_{n\to\infty}} \int_a^b (h_\varepsilon(x) + \varepsilon)\,dF_n(x) =$$

$$= \int_a^b (h_\varepsilon(x) + \varepsilon)\,dF(x) \leqslant \int_a^b g(x)\,dF(x) + 2\varepsilon$$

и аналогично

$$\underline{\lim_{n\to\infty}} \int_a^b g(x)\,dF_n(x) \geqslant \int_a^b g(x)\,dF(x) - 2\varepsilon.$$

Ввиду произвольности ε верхний и нижний пределы равны.]

13.20. Задача. Пусть $F_n \Rightarrow F$ и функция $g(x)$ непрерывна и ограничена на R. Показать, что

$$\int g(x)\,dF_n(x) \to \int g(x)\,dF(x), \quad n \to \infty.$$

[Пусть $|g(x)| \leqslant M$ при всех x. Тогда

$$\Big| \int_{\{x:\,|x| > A\}} g(x)\,dF_n(x) \Big| \leqslant M(F_n(-A) + 1 - F_n(A)),$$

$$\Big| \int_{\{x:\,|x| > A\}} g(x)\,dF(x) \Big| \leqslant M(F(-A) + 1 - F(A)).$$

Для любого $\varepsilon > 0$ выберем A настолько большим, что $F(-A) + 1 - F(A) < \varepsilon/M$ и чтобы при этом в точках $\pm A$ функция $F(x)$ была непрерывна. Используя 13.19, получаем

$$\overline{\lim_{n\to\infty}} \Big(\int_{-A}^{A} g(x)\,dF_n(x) + \int_{\{x:\,|x| > A\}} g(x)\,dF_n(x) \Big) \leqslant$$

$$\leqslant \int_{-A}^{A} g(x)\,dF(x) + \varepsilon \leqslant \int g(x)\,dF(x) + 2\varepsilon,$$

и аналогично

$$\underline{\lim_{n\to\infty}} \int g(x)\,dF_n(x) \geqslant \int g(x)\,dF(x) - 2\varepsilon.$$

Ввиду произвольности ε верхний и нижний пределы совпадают.

Из 13.20 заключаем, что если последовательность ф. р. $F_n(x)$ ($F_n(0-)=0$) слабо сходится к ф. р. $F(x)$, то соответствующая последовательность преобразований Лапласа $\psi_n(\lambda)$ сходится при всех $\lambda \geqslant 0$ к преобразованию Лапласа $\psi(\lambda)$, отвечающему ф. р. $F(x)$. Чрезвычайно важно, что справедливо обратное утверждение: из (поточечной) сходимости преобразований Лапласа вытекает слабая сходимость соответствующих функций распределения. Оба утверждения вместе называют *теоремой о непрерывности соответствия между распределениями вероятностей и их преобразованиями Лапласа*. Наиболее простое доказательство второй части теоремы непрерывности основано на следующем утверждении, которое устанавливается в 13.21—13.24:

Теорема Хелли. Пусть $F_n(x)$, $n=1, 2, \ldots$, — произвольная последовательность функций распределения. Тогда найдется подпоследовательность $F_{n_k}(x)$, $n_1 < n_2 < \ldots$, сходящаяся к некоторой неубывающей функции $F(x)$, $0 \leqslant F(x) \leqslant 1$, во всех точках ее непрерывности. Функция $F(x)$ может быть выбрана непрерывной справа; $F(x)$ удовлетворяет условиям $F(-\infty)=0$, $F(\infty)=1$, если для любого $\varepsilon>0$ можно найти такое $A>0$, что для всех n

$$F_n(-A)+1-F_n(A)<\varepsilon.$$

Последнее условие нарушается, например, для ф. р. $F_n(x)=F_n(x-n)$, где $F(x)$ — фиксированная ф. р. Очевидно, $F_n(x)\to 0$ при каждом x, так что последовательность $F_n(x)$ (и любая ее подпоследовательность) сходится к $F(x)\equiv 0$.

13.21.* З а д а ч а. Допустим, что последовательность ф. р. $F_n(x)$, $n=1, 2, \ldots$, сходится к некоторой неубывающей функции $F(x)$ при всех x из некоторого подмножества $S \subseteq R$, всюду плотного в R (т. е. для любого интервала (a, b) найдется точка $s \in S$, такая, что $s \in (a, b)$). Показать, что $F_n(x)$ сходится к $F(x)$ во всех точках непрерывности $F(x)$.

[Пусть F непрерывна в точке x; s', $s'' \in S$, $s' \in (x-\delta, x)$, $s'' \in (x, x+\delta)$. Тогда

$$F(s')=\lim_{n\to\infty} F_n(s') \leqslant \varliminf_{n\to\infty} F_n(x) \leqslant \varlimsup_{n\to\infty} F_n(x) \leqslant \lim_{n\to\infty} F_n(s'')=F(s'').$$

Выбирая здесь $\delta>0$ по $\varepsilon>0$ так, чтобы $|F(x')-F(x)|<\varepsilon$ при $|x'-x|<\delta$, получаем, что разность между верхним и нижним пределами не превосходит 2ε, а следовательно, эти пределы совпадают. Остается заметить, что $F(s') \leqslant F(x) \leqslant F(s'')$.]

13.22.* З а д а ч а. Пусть $F_n(x)$, $n=1, 2, \ldots$, — произвольная последовательность ф. р., $\{x_1, x_2, \ldots\}$ — счетное всюду плотное множество в R. С помощью *канторовского диагонального метода* построить подпоследовательность $F_{n_k}(x)$, сходящуюся при всех $x=x_i$, $i=1, 2, \ldots$.

[Рассмотрим числовую последовательность $F_n(x_1)$, $n=1, 2, \ldots$. Поскольку она ограничена, из нее можно выбрать сходящуюся подпоследовательность. Обозначим

$$n_{11}<n_{12}<\ldots<n_{1k}<\ldots$$

подпоследовательность индексов, такую, что существует предел последовательности $F_{n_{1k}}(x_1)$, $k=1, 2, \ldots$, который обозначим $F(x_1)$. Далее рассмотрим числовую последовательность $F_{n_{1k}}(x_2)$ $k=1, 2, \ldots$ и обозначим

$$n_{21}<n_{22}<\ldots<n_{2k}<\ldots$$

подпоследовательность последовательности $n_{11}, n_{12}, \ldots$, для которой сходится числовая последовательность $F_{n_{2k}}(x_2)$, $k=1, 2, \ldots$, ее предел обозначим $F(x_2)$. Продолжая этот процесс неограниченно, построим бесконечную таблицу натуральных чисел

$$n_{11}n_{12}\ldots n_{1k}\ldots$$

$$n_{21}n_{22}\ldots n_{2k}\ldots$$

$$n_{m1}n_{m2}\ldots n_{mk}\ldots$$

в которой каждая следующая строчка образована из чисел, составляющих подмножество множества чисел, из которых состоит предшествующая строчка.

Рассмотрим «диагональ»

$$n_{11}, n_{22}, \ldots, n_{mm}, \ldots,$$

Очевидно, что $n_{11}<n_{22}<\ldots<n_{mm}<\ldots$. По построению при любом m диагональная последовательность, начиная с m-го члена $(n_{mm}, n_{m+1,m+1}, \ldots)$, является подпоследовательностью каждой из первых m строчек таблицы. Это означает, что

$$\lim_{m\to\infty} F_{n_{mm}}(x_i)=F(x_i)$$

при всех $i=1, 2, \ldots$, т. е. $F_{n_{mm}}$ есть искомая подпоследовательность.]

13.23.* Задача. Пусть $F_n(x)$, $n=1, 2, \ldots$ — произвольная последовательность ф. р. Вывести из 13.21, 13.22, что существует подпоследовательность $F_{n_k}(x)$, сходящаяся к некоторой неубывающей, непрерывной справа функции $F(x)$, $0\leqslant F(x)\leqslant 1$.

[В 13.22 была построена подпоследовательность $F_{n_k}(x)$, которая сходилась во всех точках $s\in S$, где S — некоторое всюду плотное множество в R. Равенство

$$\lim_{k\to\infty} F_{n_k}(s)=F(s), \quad s\in S,$$

определяет функцию $F(s)$ на S, которая, очевидно, является неубывающей. Доопределим функцию $\overline{F}$ на всем R с сохранением монотонности, полагая для любого $x \overline{\in} S$

$$\widetilde{F}(x)=\inf\{F(s) : s>x\}, \quad \widetilde{F}(s)=F(s), \quad s\in S.$$

Итак, последовательность ф. р. $F_{n_k}(x)$ сходится на всюду плотном в R множестве S к неубывающей функции $\widetilde{F}(x)$. Ввиду 13.21 сходимость $F_{n_k}(x)\to\widetilde{F}(x)$ имеет место во всех точках непрерывности функции $\widetilde{F}(x)$. Число точек разрыва монотонной функции $\widetilde{F}(x)$ не более чем счетно. Переопределив в них функцию $\widetilde{F}(x)$ по непрерывности справа, обозначим полученную функцию $F(x)$ — она и есть искомая ф. р.]

13.24. З а д а ч а. Последовательность ф. р. $F_n(x)$ сходится к неубывающей непрерывной справа функции $F(x)$ во всех точках ее непрерывности. Допустим, что для любого $\varepsilon>0$ найдется $A>0$, такое, что

$$F_n(-A)+1-F_n(A)<\varepsilon$$

для всех n. Показать, что $F(-\infty)=0$, $F(\infty)=1$, т. е. $F(x)$ — вероятностная функция распределения.

[Число A в условии задачи можно выбрать так, чтобы в точках $\pm A$ функция $F(x)$ была непрерывна. Но тогда

$$\lim_{n\to\infty}(F_n(-A)+1-F_n(A))=F(-A)+1-F(A)\leqslant\varepsilon,$$

откуда и следует утверждение.]

Как мы видели, в качестве пределов (или частичных пределов) последовательности ф. р. $F_n(x)$ могут выступать неубывающие функции $F(x)$, для которых $F(\infty)<1$, $F(-\infty)>0$. Вся теория сходимости становится более удобной, если расширить класс рассматриваемых мер, допустив распределения на $(R, \mathscr{B})$, для которых мера всего пространства меньше единицы (или даже равна нулю). Такие распределения называем *несобственными*. Понятие слабой сходимости употребляем и тогда, когда предельная мера несобственная.

13.25. З а д а ч а. Пусть последовательность вероятностных ф. р. F_n слабо сходится к несобственной ф. р. F, $g(x)$ — непрерывна и $\lim\limits_{x\to\pm\infty} g(x)=0$. Тогда

$$\int g(x)dF_n(x)\to\int g(x)\,dF(x), \quad n\to\infty.$$

[Выводится аналогично 13.20 с использованием оценки

$$\int\limits_{\{|x|>A\}} g(x)\,dF_n(x)\leqslant\varepsilon(F_n(-A)+1-F_n(A)).]$$

13.26. Задача. Пусть $F(x)$, возможно, несобственная ф. р., $\psi(\lambda)$ — ее преобразование Лапласа. Показать, что $\psi(\lambda)$ непрерывна в нуле:

$$\int_0^\infty e^{-\lambda x}dF(x)\to\psi(0)=F(\infty)-F(0-),\quad \lambda\to 0.$$

[$$F(\infty)-F(0-)\geqslant\int_0^\infty e^{-\lambda x}dF(x)\geqslant e^{-\lambda A}(F(A)-F(0-))$$

при любом $A>0$. Устремляя λ к нулю, получаем

$$F(\infty)-F(0-)\geqslant\psi(0)\geqslant F(A)-F(0-).$$

Остается перейти к пределу при $A\to\infty$.]

13.27. Задача. Пусть последовательность вероятностных ф. р. F_n слабо сходится к несобственной ф. р. F. Тогда $\psi_n(\lambda)\to\psi(\lambda)$, $n\to\infty$, всюду, кроме точки $\lambda=0$ (здесь ψ_n и ψ — преобразования Лапласа ф. р. F_n и F соответственно, $\lambda\geqslant 0$).

[Вытекает из 13.25, 13.26.]

13.28.* Задача. Пусть последовательность преобразований Лапласа $\psi_n(\lambda)$ вероятностных ф. р. $F_n(x)$, $n=1, 2, \ldots$, сходится при всех $\lambda\geqslant 0$ к некоторому пределу $\psi(\lambda)$. Тогда $F_n\Rightarrow F$, где F, возможно, несобственная ф. р. $(F_n(0-)=0)$.

[По теореме Хелли существует подпоследовательность $F_{n_k}\Rightarrow F$, где F, возможно, несобственная ф. р. Ввиду 13.27 заключаем, что при $\lambda>0$

$$\int_0^\infty e^{-\lambda x}dF(x)=\lim_{k\to\infty}\psi_{n_k}(\lambda)=\lim_{n\to\infty}\psi_n(\lambda)=\psi(\lambda).$$

Этим условием ф. р. $F(x)$ определяется однозначно (если $\psi(\lambda)\equiv 0$, то $F(x)\equiv 0$; если $\psi(\lambda)\not\equiv 0$, то по формуле обращения после нормирования множителем $F(\infty)^{-1}$). Следовательно, все слабо сходящиеся подпоследовательности последовательности F_n имеют один и тот же предел F. Но тогда $F_n\Rightarrow F$, так как если бы в некоторой точке непрерывности x_0 ф. р. F было бы $F_n(x_0)\not\to F(x_0)$, то, выбрав подпоследовательность $F_{n_k}(x_0)\to a\neq F(x_0)$ и применив к последовательности F_{n_k} теорему Хелли, построим подпоследовательность $F_{n_{k(j)}}\Rightarrow F$, причем $F_{n_{k(j)}}(x_0)\to a\neq F(x_0)$.]

13.29. Задача. В условии 13.28 предельная ф. р. является собственной тогда и только тогда, когда $\psi(\lambda)\to 1$ при $\lambda\to 0$.

[Вытекает из 13.26.]

13.30. Задача. Пусть сл. в. $X\geqslant 0$ и при некотором n $\mathrm{M}X^n<\infty$. Тогда существует n-я производная преобразования Лапласа в точке $\lambda=0$ и $\psi^{(n)}(0)=(-1)^n\mathrm{M}X^n$.

[Для доказательства достаточно обосновать законность дифференцирования под знаком интеграла

$$\psi(\lambda)=\int_0^\infty e^{-\lambda x}dF(x)$$

в точке $\lambda=0$. Пусть для начала $n=1$. Имеем

$$\left|\int_0^\infty \frac{1-e^{-\lambda x}}{\lambda}\,dF(x)-\int_0^\infty x\,dF(x)\right| \leqslant \int_0^A \left|\frac{1-e^{-\lambda x}}{\lambda x}-1\right| x\,dF(x)+$$

$$+\int_A^\infty x\,dF(x)+\int_A^\infty \frac{1-e^{-\lambda x}}{\lambda x}\,x\,dF(x).$$

При фиксированном A перейдем к пределу при $\lambda\to 0$. Учитывая, что $(1-e^{-\lambda x})/(\lambda x)\leqslant 1$ при всех $\lambda x>0$, а при x из ограниченного множества $(1-e^{-\lambda x})/(\lambda x)\to 1$, $\lambda\to 0$, равномерно по x, получаем для модуля разности интегралов в пределе оценку сверху $2\int_A^\infty x\,dF(x)$. В предположении конечности $\mathsf{M}X$ эта оценка может быть сделана как угодно малой за счет выбора A. Таким образом, $\psi'(0)=-\mathsf{M}X$.

Точно так же показывается, что (ср. 13.16)

$$\psi'(\lambda)=-\int_0^\infty xe^{-\lambda x}dF(x),\quad \lambda>0.$$

При $n=2$ применяем аналогичные рассуждения к

$$\frac{\psi'(\lambda)-\psi'(0)}{\lambda}-\mathsf{M}X^2=\int_0^\infty\left(\frac{1-e^{-\lambda x}}{\lambda x}-1\right)x^2dF(x)$$

и т. д.]

13.31. Задача. Пусть неотрицательные сл. в. X_n, $n=1, 2, \ldots$, независимы, одинаково распределены и $\mathsf{M}X_1<\infty$. Вывести с помощью 13.28, 13.29 ЗБЧ *в форме Хинчина:*

$$n^{-1}(X_1+\ldots+X_n)\xrightarrow{\mathscr{D}}\mathsf{M}X_1.$$

[Положим $\psi(\lambda)=\mathsf{M}e^{-\lambda X_1}$. Тогда

$$\mathsf{M}e^{-\lambda n^{-1}(X_1+\ldots+X_n)}=(\psi(\lambda/n))^n.$$

Поскольку $\mathsf{M}X_1<\infty$, то (см. 13.30) существует $\psi'(0)=-\mathsf{M}X_1$. По определению производной $\psi(\lambda)=1+\psi'(0)\lambda+o(\lambda)$, $\lambda\to 0$, так что

$$(\psi(\lambda/n))^n=(1-\psi'(0)\lambda/n+o(1/n))^n\to e^{\psi'(0)\lambda}$$

при $n \to \infty$ и любом $\lambda \geqslant 0$. Следовательно, последовательность сл. в. $n^{-1}(X_1 + \ldots + X_n)$, $n=1, 2, \ldots$, сходится по распределению. Остается заметить, что $e^{-\lambda a}$ есть преобразование Лапласа распределения вероятностей, сосредоточенного целиком в точке $a \geqslant 0$.]

13.32. Задача. Пусть T_n, $n=1, 2, \ldots$, — момент первого достижения положения n в симметричном случайном блуждании (см. 11.7). Используя то, что п. ф. T_1 равна (см. 12.21)

$$Mt^{T_1} = (1 - \sqrt{1-t^2})/t, \quad 0 \leqslant t \leqslant 1,$$

показать, что последовательность T_n/n^2 сходится по распределению, применяя метод преобразований Лапласа (13.28, 13.29).
[Так как сл. в. T_n представима в виде суммы n независимых сл. в., распределенных так же, как T_1, то при $n \to \infty$

$$Me^{-\lambda T_n/n^2} \sim (1 - (1 - e^{-2\lambda/n^2})^{1/2})^n \sim (1 - \sqrt{2\lambda}/n)^n \to e^{-\sqrt{2\lambda}}.$$

Отметим, что поскольку предельная ф. р. нам известна из 11.7, то одновременно нами установлена формула

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty x^{-3/2} e^{-1/(2x)} e^{-\lambda x} dx = e^{-\sqrt{2\lambda}}.]$$

13.33.* Задача. Вывести формулу интегрирования по частям

$$\int_{a+}^{b} g(x)\, dF(x) = g(b) F(b) - g(a) F(a) - \int_a^b F(x) g'(x)\, dx,$$

предполагая, что производная $g'(x)$ непрерывна на $[a, b]$.
[Пусть $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ — произвольное разбиение отрезка $[a, b]$, $t_i \in [x_{i-1}, x_i]$, $i=1, \ldots, n$. Положим $t_0 = a$, $t_{n+1} = b$ и запишем следующую формулу *суммирования по частям*

$$\sum_{i=0}^{n} g(t_i)(F(x_i) - F(x_{i-1})) = g(b) F(b) - g(a) F(a) -$$

$$- \sum_{i=1}^{n+1} F(x_{i-1})(g(t_i) - g(t_{i-1})).$$

Устремляя диаметр разбиения к нулю, получим, что интегральная сумма в левой части написанного соотношения сходится к $\int_a^b g(x)\, dF(x)$. Используя теорему о среднем, запишем сумму в правой части в виде

$$\sum_{i=1}^{n+1} F(x_{i-1}) g'(\theta_i)(t_i - t_{i-1}), \quad \theta_i \in [t_{i-1}, t_i], \quad i=1, \ldots, n+1.$$

Эта сумма стремится к $\int\limits_a^b F(x)\,g'(x)\,dx$ (замените $g'(\theta_i)$ на $g'(x_{i-1})$ и воспользуйтесь равномерной непрерывностью функции $g'(x)$ на $[a, b]$).]

13.34. Пример (суммирование независимых величин с тяжелыми хвостами). Пусть X_i, $i=1, 2, \ldots,$ — неотрицательные, независимые, одинаково распределенные сл. в., такие, что

$$\mathbf{P}(\omega : X_1(\omega) > x) = 1 - F(x) \sim ax^{-\alpha}, \quad x \to \infty,$$

где $0<\alpha<1$. В таком случае $\mathbf{M}X_1=\infty$ (см. 7.33) и мы покажем, что в сумму $S_n=X_1+\ldots+X_n$ существенный вклад вносит максимальное слагаемое $M_n=\max(S_1, \ldots, S_n)$ (ср. 11.7).

Сл. в. $M_n/n^{1/\alpha}$ имеют невырожденное предельное распределение:

$$\mathbf{P}(\omega : M_n(\omega)/n^{1/\alpha} \leqslant x) = F(xn^{1/\alpha})^n = (1-(1-F(xn^{1/\alpha})))^n =$$

$$= (1 - ax^{-\alpha}n^{-1}(1+o(1)))^n \to e^{-ax^{-\alpha}}, \quad x \geqslant 0, \quad n \to \infty.$$

Покажем, что также сл. в. $S_n/n^{1/\alpha}$ имеют невырожденное предельное распределение. С частным случаем $\alpha=1/2$ такого рода мы встречались в 11.7, однако там доказательство целиком опиралось на конкретный вид функции распределения сл. в. $S_n/n^{1/\alpha}$, тогда как здесь получается результат весьма общего характера. Примечательно, что скорость «расползания» с ростом n вероятностной массы с законом распределения сл. в. S_n, выражающаяся в множителе $n^{-1/\alpha}$, и сам предельный закон для $S_n n^{-1/\alpha}$ не зависят ни от каких частных особенностей ф. р. $F(x)$, кроме как от асимптотики ее хвостов. Доказательство предельной теоремы основано на применении теоремы непрерывности для преобразований Лапласа. Запишем

$$\mathbf{M}e^{-\lambda S_n n^{-1/\alpha}} = \psi(\lambda n^{-1/\alpha})^n, \quad \psi(\lambda) = \mathbf{M}e^{-\lambda X_1}.$$

Получим асимптотику для $1-\psi(\lambda)$ при $\lambda \to 0$. Интегрируя по частям, запишем

$$\int\limits_0^A e^{-\lambda x}dF(x) = e^{-\lambda x}F(x)\,\Big|_{0-}^{A} + \int\limits_0^A \lambda e^{-\lambda x}F(x)\,dx.$$

Устремляя A к ∞, имеем

$$\psi(\lambda) = \lambda\int\limits_0^\infty e^{-\lambda x}F(x)\,dx \text{ или } 1-\psi(\lambda) = \lambda\int\limits_0^\infty e^{-\lambda x}(1-F(x))\,dx.$$

Покажем, что

$$\int\limits_0^\infty e^{-\lambda x}(1-F(x))\,dx \sim \int\limits_0^\infty e^{-\lambda x}\,ax^{-\alpha}\,dx, \quad \lambda \to 0.$$

По любому $\varepsilon > 0$ выберем $A > 0$ таким, что

$$(a-\varepsilon)x^{-\alpha} < 1 - F(x) < (a+\varepsilon)x^{-\alpha} \text{ при } x > A.$$

Тогда

$$\int_A^\infty e^{-\lambda x}(1-F(x))\,dx \lesseqgtr (a \pm \varepsilon)\int_A^\infty e^{-\lambda x}x^{-\alpha}dx =$$

$$=(a\pm\varepsilon)\int_{\lambda A}^\infty e^{-y}y^{-\alpha}dy\cdot\lambda^{-1+\alpha} \sim (a\pm\varepsilon)\Gamma(1+\alpha)\lambda^{-1+\alpha}, \quad \lambda \to 0,$$

$$\int_0^A e^{-\lambda x}(1-F(x))\,dx \leqslant A.$$

Следовательно,

$$\int_0^\infty e^{-\lambda x}(1-F(x))\,dx \sim a\Gamma(1+\alpha)\lambda^{-1+\alpha}, \quad \lambda \to 0,$$

$$1-\psi(\lambda) \sim a\Gamma(1+\alpha)\lambda^{\alpha}, \quad \lambda \to 0.$$

Подставляя полученную асимптотику в выражение для преобразования Лапласа сл. в. $S_n n^{-1/\alpha}$, имеем

$$\psi(\lambda n^{-1/\alpha})^n = (1-(1-\psi(\lambda n^{-1/\alpha})))^n =$$
$$=(1-a\Gamma(1+\alpha)n^{-1}+o(n^{-1}))^n \to e^{-a\Gamma(1+\alpha)\lambda^\alpha}$$

Применение утверждения 13.28 устанавливает сходимость последовательности $S_n n^{-1/\alpha}$ по распределению к сл. в. с преобразованием Лапласа $\exp(-a\Gamma(1+\alpha)\lambda^\alpha)$.

В преобразовании Лапласа

$$\psi_c(\lambda) = \exp(-(c\lambda)^\alpha)$$

параметр c является масштабным. Так как

$$\psi_{c_1}(\lambda)\,\psi_{c_2}(\lambda) = \psi_{(c_1^\alpha+c_2^\alpha)^{1/\alpha}}(\lambda),$$

то сумма двух независимых сл. в. X_1 и X_2 с преобразованиями Лапласа $\psi_{c_1}(\lambda)$ и $\psi_{c_2}(\lambda)$ имеет распределение вероятностей, отличающееся от распределения слагаемых лишь масштабным множителем. Это свойство сохранения *типа распределения* при суммировании независимых слагаемых называется *устойчивостью.*

В случае n независимых одинаково распределенных слагаемых с преобразованием Лапласа $\psi_1(\lambda)$ получаем, что сумма $X_1+\ldots+X_n$ распределена как $n^{1/\alpha}X_1$. Сл. в. с одинаковыми распределениями называют *совпадающими* или *равными по распределению* и между ними ставят знак $\overset{d}{=}$. Таким образом,

$$(X_1+\ldots+X_n)/n^{1/\alpha}\overset{d}{=}X_1,$$

т. е. для нарастающих сумм устойчивых сл. в. предельная теорема заменяется точным равенством. Другой хорошо нам известный пример устойчивого распределения — нормальное распределение с нулевым средним.

13.35. Задача. Методом преобразования Лапласа найти предельное распределение для геометрического $(p(1-p)^k,\ k=0,\ 1,\ 2,\ \ldots)$ и отрицательного биномиального $(C^k_{r+k-1}p^r(1-p)^k,\ k=0,\ 1,\ 2,\ \ldots)$ распределений при $p\to 0$ (после надлежащей нормировки).

[Распределение вероятностей сл. в. $X=X_p$ с геометрическим распределением $p(1-p)^k$, $k=0,\ 1,\ 2,\ \ldots$, при $p\to 0$, очевидно, «расползается» по положительной полупрямой. Поэтому требуется нормирующий множитель $a=a_p\to 0$, который необходимо выбрать из условия сходимости по распределению сл. в. a_pX_p при $p\to 0$. Преобразование Лапласа

$$Me^{-\lambda aX}=\sum_{k=0}^{\infty}e^{-\lambda ak}p(1-p)^k=p(1-(1-p)e^{-\lambda a})^{-1}=$$

$$=p(1-e^{-\lambda a}+pe^{-\lambda a})^{-1}=p(\lambda a+p+o(a)+o(p))^{-1}$$

стремится к $(1+\lambda)^{-1}$, если взять $a=p$. Предельное распределение является экспоненциальным (см. 13.15). Преобразование Лапласа отрицательного биномиального распределения — r-я степень геометрического, так что предельным здесь будет гамма-распределение (см. 13.15, ср. 3.14).]

Преобразование Лапласа может быть определено и для распределений на всей прямой:

$$\int_{-\infty}^{\infty}e^{-\lambda x}\,dF(x)$$

при условии, что интеграл конечен на каком-нибудь интервале значений λ. Однако для распределений на всей прямой такого рода условие является слишком ограничительным. Для них вместо преобразования Лапласа рассматривают *характеристическую функцию* (х. ф.)

$$\psi_X(t)=Me^{itX}\equiv M\cos tX+iM\sin tX,\quad -\infty<t<\infty.$$

Нетрудно видеть, что

$$\psi_X(t)=\int_{-\infty}^{\infty}\cos tx\,dF(x)+i\int_{-\infty}^{\infty}\sin tx\,dF(x)\equiv\int_{-\infty}^{\infty}e^{itx}\,dF(x).$$

В самом деле, если $X^{(n)}$ — последовательность сл. в., участвующих в определении $\mathbf{M}X$, то (см. 7.19)

$$\mathbf{M}\cos tX=\lim_{n\to\infty}\mathbf{M}\cos tX^{(n)}=$$

$$=\lim_{n\to\infty}\sum_{k=-\infty}^{\infty}\cos(tkn^{-1})(F((k+1)n^{-1})-F(kn^{-1}))=\int_{-\infty}^{\infty}\cos tx dF(x)$$

и аналогичное соотношение верно для $\mathbf{M}\sin tX$. Если $F(x)$ дифференциируема, то характеристическая функция

$$\psi(t)=\int_{-\infty}^{\infty}e^{itx}f(x)\,dx,\quad f(x)=F'(x),$$

— обычное *преобразование Фурье* плотности $f(x)$.

Отметим некоторые простые свойства характеристических функций. Во-первых, х. ф. $\psi_X(t)$ определена для любой сл. в. X, так как математическое ожидание ограниченной сл. в. существует и конечно. Так как $|e^{itX}|\leqslant 1$, то $|\psi_X(t)|\leqslant 1$. Если X_1, X_2 независимы, то

$$\psi_{X_1+X_2}(t)=\psi_{X_1}(t)\,\psi_{X_2}(t).$$

Для любых постоянных a, b $\psi_{aX+b}(t)=e^{ib}\psi_X(at)$. В задачах 13.36, 13.37 устанавливается, что х. ф. однозначно определяет распределение вероятностей.

13.36. Задача. Рассмотрим семейство функций $X^{\delta}_{a,b}(x)$, $a<b$, $\delta>0$, полагая $X^{\delta}_{a,b}(x)=1$ при $x\in[a, b]$, $X^{\delta}_{a,b}(x)=0$ при $x\overline{\in}[a-\delta, b+\delta]$ и определяя $X^{\delta}_{a,b}(x)$ на интервалах $(a-\delta, a)$, $(b, b+\delta)$ любым образом, но так, чтобы функция $X^{\delta}_{a,b}(x)$ была непрерывно дифференцируема и $0\leqslant X^{\delta}_{a,b}(x)\leqslant 1$. Показать, что совокупность значений интегралов

$$\int_{-\infty}^{\infty}X^{\delta}_{a,b}(x)\,dF(x)$$

при всевозможных a, b, δ однозначно определяет распределение F. [Очевидно, что

$$F(b)-F(a)\leqslant\int X^{\delta}_{a,b}(x)\,dF(x)\leqslant F(b+\delta)-F(a-\delta).$$

Выбирая в качестве $a<b$ любые точки непрерывности ф. р. и устремляя δ к 0, получаем значение $F(b)-F(a)$ как предел соответствующих интегралов.]

13.37.* Задача. Используя теорему о разложении непрерывно дифференцируемой функции в ряд Фурье, показать, что значения интегралов

$$\psi(t)=\int_{-\infty}^{\infty} e^{itx}\,dF(x)$$

при всевозможных действительных t однозначно определяют распределение F (другое доказательство см. 13.54).

[Достаточно показать, что для любой функции $X^{\delta}_{a,b}(x)$ из 13.36 и любого $\varepsilon>0$ найдется линейная комбинация значений характеристической функции (построенная без знания $F(x)$)

$$\sum_{\{k\}} c_k\psi(t_k)=\int_{-\infty}^{\infty}\sum_{\{k\}} c_k e^{it_kx}\,dF(x),$$

такая, что

$$\left|\sum_{\{k\}} c_k\psi(t_k)-\int_{-\infty}^{\infty} X^{\delta}_{a,b}(x)\,dF(x)\right|<\varepsilon.$$

Выберем N таким, чтобы $[a, b]\subseteq[-N, N]$ и

$$F(-N)+1-F(N)<\varepsilon/2.$$

Функцию $X^{\delta}_{a,b}(x)$ продолжим периодически за пределы отрезка $[-N, N]$, и, разложив ее в равномерно сходящийся ряд Фурье, возьмем его отрезок $h(x)=\sum_{\{k\}} c_k e^{it_kx}$ так, чтобы

$$|h(x)-X^{\delta}_{a,b}(x)|<\varepsilon/2,\quad x\in[-N, N].$$

Отсюда получаем

$$\int |h(x)-X^{\delta}_{a,b}(x)|\,dF(x)\leqslant\frac{\varepsilon}{2}\int_{-N}^{N} dF(x)+F(-N)+1-F(N)<\varepsilon\Bigg].$$

13.38. Пример (характеристическая функция нормального распределения). Вычислим интеграл

$$\mathcal{J}(t)=\frac{1}{2}\int_{-\infty}^{\infty} e^{itx}e^{-x^2}\,dx=\int_{0}^{\infty}\cos tx\,e^{-x^2}\,dx.$$

[Интегрируя по частям, найдем

$$\mathcal{J}(t)=t^{-1}\int_{0}^{\infty} e^{-x^2}d\sin tx=t^{-1}e^{-x^2}\sin tx\,\Big|_0^{\infty}+$$

$$+t^{-1}\int_{0}^{\infty}\sin tx\cdot 2xe^{-x^2}\,dx=2t^{-1}\int_{0}^{\infty} x\sin tx\cdot e^{-x^2}\,dx.$$

С другой стороны, дифференцируя под знаком интеграла, находим

$$\mathcal{I}'(t) = -\int_0^\infty x \sin tx \cdot e^{-x^2}\,dx.$$

Таким образом, для $\mathcal{I}(t)$ получаем дифференциальное уравнение

$$\mathcal{I}'(t) = -(1/2)\,t\mathcal{I}(t)$$

с начальным условием $\mathcal{I}(0) = \int_0^\infty e^{-x^2}\,dx = \sqrt{\pi}/2$. Поделив на $\mathcal{I}(t)$, имеем

$$\mathcal{I}'(t)/\mathcal{I}(t) = -(1/2)\,t \text{ или } (\ln \mathcal{I}(t))' = -(1/2)\,t,$$

откуда

$$\ln \mathcal{I}(t) = -(1/4)\,t^2 + c, \quad \mathcal{I}(t) = \exp(-t^2/4)\cdot\sqrt{\pi}/2.$$

Характеристическая функция стандартного нормального распределения:

$$\psi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-x^2/2}\,dx = \frac{2}{\sqrt{\pi}}\,\mathcal{I}\left(t\sqrt{2}\right) = e^{-t^2/2}.\Big]$$

13.39. Задача. Пусть $F(x)$, возможно, несобственная ф.р., $\psi(t)$ — ее характеристическая функция. Показать, что $\psi(t)$ непрерывна в нуле:

$$\psi(t) = \int_{-\infty}^{\infty} e^{itx}\,dF(x) \to F(\infty) - F(-\infty), \quad t \to 0.$$

$$\Big[\,|\psi(t) - (F(\infty) - F(-\infty))| \leqslant \int_{-\infty}^{\infty} |e^{itx} - 1|\,dF(x) \leqslant$$

$$\leqslant \int_{-A}^{A} |e^{itx} - 1|\,dF(x) + 2\,(F(-A) + 1 - F(A)).$$

Устремляя t к нулю, а затем A к ∞, получаем требуемое.]

13.40. Задача. Доказать *теорему непрерывности* для характеристических функций: если $F_n \Rightarrow F$, то $\psi_n(t) \to \psi(t)$ при всех $t \neq 0$; обратно, если последовательность х. ф. $\psi_n(t)$, $n = 1, 2, \ldots$, сходится при всех t к некоторому пределу $\psi(t)$, то $F_n \Rightarrow F$, где F — возможно несобственная ф.р. (т. е. $F(-\infty) > 0$ или $F(\infty) < 1$, или и то и другое). Предельная ф.р. — собственная тогда и только тогда, когда $\psi(t) \to 1$ при $t \to 0$.
[Устанавливается аналогично 13.27—13.29.]

13.41. Задача. Показать, что характеристическая функция равномерно непрерывна.

$$\left[|\psi(t_1)-\psi(t_2)| \leqslant \int |e^{it_1x}-e^{it_2x}|\,dF(x) = \int |e^{i(t_1-t_2)x}-1|\,dF(x), \right.$$

далее, как в 13.39.]

13.42. Задача. Если $\mathrm{M}|X|^n$ конечно при некотором натуральном n, то существует n-я производная $\psi^{(n)}(t)$ и $\psi^{(n)}(0) = (it)^n \mathrm{M}X^n$. Производная $\psi^{(n)}(t)$ равномерно непрерывна. [Устанавливается аналогично 13.30.]

13.43. Задача. Сл. в. X_n, $n=1, 2, \ldots$, независимы, одинаково распределены, $\mathrm{M}|X_1| < \infty$. Показать, что $n^{-1}(X_1 + \ldots + X_n) \xrightarrow{\mathscr{D}} \mathrm{M}X_1$. [Устанавливается аналогично 13.31.]

13.44. Задача. Доказать, применяя теорему непрерывности 13.40, *центральную предельную теорему* для независимых одинаково распределенных слагаемых: пусть X_n, $n=1, 2, \ldots$, независимы, одинаково распределены, имеют конечные $\mathrm{M}X_1$, $\mathrm{D}X_1$. Показать, что последовательность центрированных и нормированных сумм $(X_1 + \ldots + X_n - n\mathrm{M}X_1)/\sqrt{n\mathrm{D}X_1}$ сходится по распределению к стандартной нормальной величине.

[Положим $Y_i = (X_i - \mathrm{M}X_i)/\sqrt{\mathrm{D}X_i}$, $i=1, 2, \ldots$, так что

$$(X_1 + \ldots + X_n - n\mathrm{M}X_1)/\sqrt{n\mathrm{D}X_1} = (Y_1 + \ldots + Y_n)/\sqrt{n}.$$

Поскольку $\mathrm{M}Y_k = 0$, $\mathrm{D}Y_k = \mathrm{M}Y_k^2 = 1$, то (см. 13.42)

$$\mathrm{M}e^{itY_1} = 1 - t^2/2 + o(t^2),$$

$$\mathrm{M}\exp\left(it(Y_1 + \ldots + Y_n)/\sqrt{n}\right) = (1 - t^2/(2n) + o(1/n))^n,$$

что стремится при $n \to \infty$ к $e^{-t^2/2}$ — характеристической функции стандартного нормального распределения.]

Записав утверждение 13.44 в виде

$$F_{X_1+\ldots+X_n}\left(na + x\sigma\sqrt{n}\right) \to \Phi(x), \quad n \to \infty, \quad a = \mathrm{M}X_1, \quad \sigma^2 = \mathrm{D}X_1,$$

применим его для конечных, но «больших» значений n, заменив знак стремления на приближенное равенство:

$$F_{X_1+\ldots+X_n}\left(na + x\sigma\sqrt{n}\right) \approx \Phi(x),$$

или

$$F_{X_1+\ldots+X_n}(y) \approx \Phi\left((y - na)/(\sigma\sqrt{n})\right).$$

Прежде всего отметим важность этой формулы для вычислений. Нахождение функции распределения суммы большого числа независимых слагаемых в явном виде в общем случае представляет собой неразрешимую задачу. Использование нормального приближения, как показывает опыт, дает хорошие результаты даже при небольших n. Еще более важно то обстоятельство, что распределение вероятностей суммы $X_1+\ldots+X_n$ при больших n определяется (приближенно) лишь значениями $\mathbf{M}X_1$, $\mathbf{D}X_1$ и не зависит (с точностью до приближения, связанного с конечностью n) от частного вида ф. р. $F(x)=\mathbf{P}(\omega:X_1(\omega)\leqslant x)$.

Результат 13.44 допускает далеко идущие обобщения. Во-первых, условие независимости сл. в. $X_1, X_2, \ldots$ не является необходимым, его можно заменить некоторыми условиями слабой зависимости. Исследование условий зависимости, при которых удается доказать центральную предельную теорему для сумм случайных величин, составляет обширный раздел современной теории вероятностей. Во-вторых, сл. в. $X_1, X_2, \ldots$ можно взять разнораспределенными (и независимыми). Посмотрим, как нужно трансформировать для этого случая доказательство теоремы 13.44 и каковы должны быть ограничения на сл. в. $X_1, X_2, \ldots$.

Итак, пусть X_k, $k=1, 2, \ldots,$ — независимые случайные величины с ф. р. $F_k(x)=\mathbf{P}(\omega:X_k(\omega)\leqslant x)$ и конечными дисперсиями $\sigma^2_k=\mathbf{D}X_k$. Не ограничивая общности, предположим, что $\mathbf{M}X_k=0$, $\sigma^2_k>0$, $k=1, 2, \ldots$. Полагая $S_n=X_1+\ldots+X_n$, $s_n=(\sigma_1^2+\ldots+\sigma^2_n)^{1/2}$ ($s_n=(n\mathbf{D}X_1)^{1/2}$ в случае одинаково распределенных слагаемых), посмотрим, как можно обеспечить справедливость утверждения

$$F_{S_n/s_n}(x)\to\Phi(x),\quad n\to\infty.$$

Имеем

$$\mathbf{M}e^{itS_n/s_n}=\prod_{k=1}^{n}\mathbf{M}e^{i(t/s_n)X_k}.$$

В предположении, что $s_n\to\infty$ при $n\to\infty$

$$\mathbf{M}e^{i(t/s_n)X_k}=1-(1/2)\,t^2\sigma_k^2/s_n^2+r_{nk},$$

где $r_{nk}=o(1/s^2_n)$ при $n\to\infty$ и фиксированных t и k. Если бы остатками r_{nk}, $k=1,\ldots,n$, можно было пренебречь и если выполняется условие

$$\text{(A)}\quad \max_{(k\leqslant n)}\sigma_k^2/s_n^2\to 0,\quad n\to\infty,$$

то

$$\mathbf{M}e^{itS_n/s_n}\sim\prod_{k=1}^{n}\left(1-\frac{1}{2}\frac{t^2\sigma_k^2}{s_n^2}\right)=\exp\left(\sum_{k=1}^{n}\ln\left(1-\frac{1}{2}\frac{t^2\sigma_k^2}{s_n^2}\right)\right)=$$

$$=\exp\left(\sum_{k=1}^{n}\left(-\frac{1}{2}\frac{t^2\sigma_k^2}{s_n^2}+o(1)\frac{\sigma_k^2}{s_n^2}\right)\right)=\exp\left(-\frac{t^2}{2}+o(1)\right)$$

(отметим, что множители $o(1)$ под знаком суммы малы равномерно по $k\leqslant n$, а переход к логарифмам законен, если n достаточно велико). Требование (A) равномерной малости дисперсий $\sigma_k{}^2$, $k=$

$=1, \ldots, n$, слагаемых $X_1, \ldots, X_n$ относительно дисперсии $s^2{}_n$ суммы $S_n = X_1 + \ldots + X_n$ является, как видно, необходимым для применения схемы доказательства 13.44 (в условиях 13.44 оно выполняется автоматически, так как $\sigma^2{}_k = \sigma^2 = \mathrm{const}$).

Из проведенных выкладок легко усмотреть, что следующее требование

$$\text{(B)}\quad \sum_{k=1}^{n} |r_{nk}| \to 0, \quad n \to \infty$$

позволяет пренебречь остатками r_{nk}, $k=1, \ldots, n$, так что при выполнении (A), (B) имеем

$$\mathbf{M} e^{itS_n/s_n} \to e^{-t^2/2}, \quad n \to \infty,$$

т. е. последовательность сл. в. S_n/s_n, $n=1, 2, \ldots$, сходится по распределению к стандартной нормальной величине. В нижеследующих задачах 13.47, 13.50, 13.52 приведены условия, обеспечивающие выполнение требования (B), выражающиеся явным образом через ф. р. F_k, $k=1, 2, \ldots$.

13.45. Задача. Пусть сл. в. X имеет конечное математическое ожидание, функция $g(x)$ — выпукла вниз, т. е. для любого a существует такое λ, что

$$g(x) \geqslant g(a) + \lambda(x-a)$$

(для всех x, a из области определения функции $g(x)$). Вывести *неравенство Иенсена* (выпуклых функций)

$$\mathbf{M} g(X) \geqslant g(\mathbf{M} X).$$

[Взять математическое ожидание от обеих частей неравенства

$$g(X) \geqslant g(\mathbf{M}X) + \lambda(X - \mathbf{M}X).]$$

13.46. Задача. Пусть $0 < \alpha < \beta$, сл. в. X неотрицательна. Показать, что

$$(\mathbf{M}X^\alpha)^{1/\alpha} \leqslant (\mathbf{M}X^\beta)^{1/\beta}.$$

[Если $\mathbf{M}X^\beta = \infty$, то неравенство тривиально. При $\mathbf{M}X^\beta < \infty$ также и $\mathbf{M}X^\alpha < \infty$ (см. 7.34) функция $g(x) = x^{\beta/\alpha}$ выпукла вниз, так что, применяя неравенство Иенсена к сл. в. X^α, получаем

$$(\mathbf{M}X^\alpha)^{\beta/\alpha} = g(\mathbf{M}X^\alpha) \leqslant \mathbf{M} g(X^\alpha) = \mathbf{M}X^\beta.]$$

13.47. Задача. Сл. в. X_k, $k=1, 2, \ldots$, независимы, $\mathbf{M}X_k = 0$,

$$\mathbf{M}X_k^2 = \sigma_k^2, \quad 0 < \sigma_k^2 < \infty, \quad \mathbf{M}|X_k|^3 < \infty, \quad s_n = (\sigma_1^2 + \ldots + \sigma_n^2)^{1/2}.$$

Предположим, что выполнено *условие Ляпунова*

$$\text{(C)}\quad \sum_{k=1}^{n} \mathbf{M}|X_k|^3/s_n^3 \to 0, \quad n \to \infty.$$

Показать, что тогда выполнено условие

$$\text{(A)} \quad \max_{(k\leqslant n)} \sigma_k^2/s_n^2 \to 0, \quad n\to\infty.$$

[Из 13.46 получаем

$$\sigma_k^3 = (\mathbf{M}X_k^2)^{3/2} \leqslant \mathbf{M}|X_k|^3,$$

так что

$$\max_{(k\leqslant n)} \sigma_k^3/s_n^3 \leqslant \sum_{k=1}^{n} \sigma_k^3/s_n^3 \to 0, \quad n\to\infty.]$$

13.48.* Задача. Установить неравенство

$$\left| e^{it} - \sum_{k=0}^{m-1} \frac{(it)^k}{k!} \right| \leqslant \frac{|t|^m}{m!}.$$

$\left[|e^{it}-1| = \left| \int\limits_0^t e^{is}\,ds \right| \leqslant |t|\right.$. Далее по индукции

$$\left| e^{it} - \sum_{k=0}^{m} \frac{(it)^k}{k!} \right| = \left| \int\limits_0^t \left(e^{is} - \sum_{k=1}^{m} \frac{(is)^{k-1}}{(k-1)!} \right) ds \right| =$$

$$= \left| \int\limits_0^t \left(e^{is} - \sum_{k=0}^{m-1} \frac{(is)^k}{k!} \right) ds \right| \leqslant \int\limits_0^{|t|} \frac{|s|^m}{m!}\, ds = \frac{|t|^{m+1}}{(m+1)!}.\Bigg]$$

13.49. Задача. Вывести из 13.48 следующее неравенство для характеристической функции $\psi_X(t) = \mathbf{M}\exp(itX)$:

$$\left| \psi_X(t) - \sum_{k=0}^{m-1} \psi_X^{(k)}(0) \frac{t^k}{k!} \right| \leqslant \mathbf{M}|X|^m \frac{|t|^m}{m!},$$

где предполагается, что $\mathbf{M}|X|^m < \infty$ $(\psi^{(k)}(t) = d^k\psi(t)/dt^k)$.

[Для левой части неравенства 13.49 имеем с учетом 13.48

$$\left| \int\limits_{-\infty}^{\infty} e^{itx}\, dF_X(x) - \sum_{k=0}^{m-1} \frac{t^k}{k!} \int\limits_{-\infty}^{\infty} (ix)^k\, dF_X(x) \right| \leqslant$$

$$\leqslant \int\limits_{-\infty}^{\infty} \left| e^{itx} - \sum_{k=0}^{m-1} \frac{(itx)^k}{k!} \right| dF_X(x) \leqslant \int\limits_{-\infty}^{\infty} \frac{|tx|^m}{m!}\, dF_X(x) \leqslant \mathbf{M}|X|^m \frac{|t|^m}{m!}.\Bigg]$$

13.50. Задача. В условиях 13.47 показать, что

$$\text{(B)} \quad \sum_{k=1}^{n} |r_{nk}| \to 0, \quad n\to\infty,$$

где

$$r_{nk}=Me^{i(t/s_n)X_k}-1+(1/2)\,t^2\sigma_k^2/s_n^2.$$

[Учитывая, что

$$0=MX_k=\int_{-\infty}^{\infty} x\,dF_k(x),\ \ \sigma_k^2=\int_{-\infty}^{\infty} x^2\,dF_k(x),$$

запишем

$$r_{nk}=\int_{-\infty}^{\infty}\left(e^{i(t/s_n)x}-1-\frac{itx}{s_n}+\frac{1}{2}\,\frac{t^2x^2}{s_n^2}\right)dF_k(x).$$

Применяя неравенство 13.48 при $m=3$ и суммируя по k, получаем

$$\sum_{k=1}^{n}|r_{nk}|\leqslant\sum_{k=1}^{n}\int_{-\infty}^{\infty}\frac{|tx|^3}{s_n^3}\,dF_k(x)=\frac{|t|^3}{s_n^3}\sum_{k=1}^{n}M|X_k|^3,$$

что стремится к нулю при $n\to\infty$ ввиду условия Ляпунова 13.47 (С).]

13.51. Задача. Доказать теорему Ляпунова: если сл. в. X_1, $X_2,\ldots$ независимы, $MX_k=0$, $MX_k{}^2=\sigma^2{}_k$, $0<\sigma^2{}_k<\infty$ и выполнено условие

$$(\mathrm{C})\quad \sum_{k=1}^{n} M|X_k|^3/s_n^3\to 0,\quad n\to\infty,$$

то предельное распределение последовательности сл. в. S_n/s_n, $S_n=X_1+\ldots+X_n$, $s^2{}_n=\sigma_1{}^2+\ldots+\sigma^2{}_n$, $n=1,\ 2,\ldots,$ — стандартное нормальное.

Вытекает из 13.47, 13.50 и рассуждений, предшествующих задаче 13.45.]

13.52.* Задача. Заменив в 13.51 условие (С) на *условие Линдеберга*

$$(\mathrm{D})\quad \frac{1}{s_n^2}\sum_{k=1}^{n}\ \int\limits_{\{|x|>\varepsilon s_n\}} x^2\,dF_k(x)\to 0,\quad n\to\infty,$$

при любом $\varepsilon>0$, показать, что выполняются соотношения 13.47 (А) и 13.50 (В) и, следовательно, предельное распределение последовательности сл. в. S_n/s_n — стандартное нормальное.

$$\left[\sigma_k^2=\int\limits_{\{|x|\leqslant\varepsilon s_n\}} x^2\,dF_k(x)+\int\limits_{\{|x|>\varepsilon s_n\}} x^2\,dF_k(x)\leqslant\varepsilon^2 s_n^2+\int\limits_{\{|x|>\varepsilon s_n\}} x^2\,dF_k(x).\right.$$

Учитывая, что максимум из положительных чисел не превосходит их суммы, получаем отсюда

$$\max_{\{k\leqslant n\}}\sigma_k^2/s_n^2\leqslant\varepsilon^2+\sum_{k=1}^{n}\ \int\limits_{\{|x|>\varepsilon s_n\}} x^2\,dF_k(x),$$

откуда следует соотношение 13.47 (А).

В представлении (см. 13.50).

$$r_{nk}=\int_{-\infty}^{\infty}\left(e^{i(t/s_n)x}-1-\frac{itx}{s_n}+\frac{1}{2}\frac{t^2x^2}{s_n^2}\right)dF_k(x)$$

разобьем область интегрирования на части $|x|\leqslant\varepsilon s_n$ и $|x|>\varepsilon s_n$ и заменим модуль подынтегральной функции соответственно на (см. 13.48 при $m=3$, 2).

$$(1/3!)\,|tx|^3/s_n^3\leqslant(1/3!)\,\varepsilon|t|^3x^2/s_n^2$$

и на

$$\left|e^{itx/s_n}-1-\frac{itx}{s_n}\right|+\frac{t^2x^2}{2s_n^2}\leqslant\frac{t^2x^2}{s_n^2}.$$

Суммируя по k, получаем

$$\sum_{k=1}^{n}|r_{nk}|\leqslant\frac{\varepsilon|t|^3}{3!}+\frac{t^2}{s_n^2}\sum_{k=1}^{n}\int\limits_{\{|x|>\varepsilon s_n\}}x^2\,dF_k(x),$$

откуда следует соотношение 13.50 (B).]

13.53. Задача. Показать, что условие Линдеберга 13.52 (D) выполняется в предположениях задач 13.44 и 13.51.

[В случае 13.44 условие (D) сводится к соотношению

$$\int\limits_{\{|x|>\varepsilon\sigma\sqrt{n}\}}x^2\,dF(x)\to 0,\quad n\to\infty,$$

а в случае 13.51 оно получается из оценки

$$\int\limits_{\{|x|>\varepsilon s_n\}}x^2\,dF_k(x)\leqslant\frac{1}{\varepsilon s_n}\int\limits_{\{|x|>\varepsilon s_n\}}|x|^3\,dF_k(x).\Bigg]$$

13.54. Задача. Воспользовавшись тем, что $N\sin(x/N)\to x$ при $N\to\infty$ равномерно на любом отрезке, вывести из теоремы Вейерштрасса о приближении непрерывной функции многочленами (см. 8.11), что непрерывная на отрезке функция может быть равномерно приближена функциями вида

$$h(x)=\sum_{\{k\}}c_ke^{it_kx},$$

где t_k — действительные числа, k пробегает конечное множество. Опираясь на этот результат, провести заново доказательство утверждения 13.37.

[Ввиду теоремы Вейерштрасса достаточно показать, что функциями $h(x)$ можно приблизить равномерно многочлены, а для этого, в свою очередь, достаточно проверить, что любой одночлен x^n,

$n=1, 2, \ldots$, приближается равномерно на любом отрезке функциями $h(x)$. Но

$$N^n \sin(x/N)^n \to x^n, \quad N \to \infty,$$

и остается заметить, что

$$\sin x = (e^{ix} - e^{-ix})/2.$$

В доказательстве утверждения 13.37 ссылку на свойство рядов Фурье надо заменить ссылкой на только что установленный результат.]

13.55. Задача. Вывести следующую формулу обращения для целочисленной сл. в. X:

$$p_X(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-int} \psi_X(t)\, dt, \quad \psi_X(t) = \mathbf{M} e^{itX}$$

$$\Big[\int_{-\pi}^{\pi} e^{-int} \psi_X(t)\, dt = \int_{-\pi}^{\pi} \sum_{(k)} p_X(k)\, e^{it(k-n)}\, dt =$$

$$= \sum_{(k)} p_X(k) \int_{-\pi}^{\pi} e^{it(k-n)}\, dt = p_X(n) \cdot 2\pi,$$

где почленное интегрирование законно ввиду равномерной сходимости ряда.]

§ 14. МАРКОВСКИЕ МОДЕЛИ

Многие модели процессов, с которыми мы познакомились выше, обладали *марковским свойством*, или свойством *отсутствия последействия*: при фиксированном «настоящем» «прошлое» и «будущее» процесса (условно) независимы. Процессы, обладающие этим свойством, были названы марковскими. Эволюцию марковского процесса удобно представлять себе в виде траектории в «фазовом пространстве» $\mathscr{X}$ — общем множестве значений сл. в. $X_t(\omega)$, $t \in T$. Весьма часто говорят о «системе», состояние которой описывается точкой $x \in \mathscr{X}$, и интерпретируют $X_t(\omega)$ как состояние системы в момент времени t. Тот факт, что процесс изменения состояний начиная с любого момента времени t зависит лишь от X_t, или, иначе, вся зависимость от предыстории X_s, $s \leqslant t$, заключена в значении X_t, делает описание марковских процессов относительно простым: необходимо знать лишь «локальные» характеристики движения при сдвиге на бесконечно малое время, если t непрерывно, и на один шаг, если время дискретно. Начнем с наиболее простого случая, когда $\mathscr{X}$ — конечное или счетное множество, $t=0, 1, 2, \ldots$. Это так называемые марковские цепи с дискретными множеством состояний и временем. Перенумеровав множество $\mathscr{X}$, удобно в этом случае принять в качестве фазового пространства множество номеров (подмножество целых чисел).

Предположим, что последовательность сл. в. X_n, $n=0, 1, 2, \ldots$, образует марковскую цепь с некоторым множеством целых чисел $\mathscr{I}$ в качестве пространства состояний. Это означает, что условные вероятности

$$p_{X_n|X_0, X_1, \ldots, X_{n-1}}(i_n | i_0, i_1, \ldots, i_{n-1})$$

при любых n, $i_0, i_1, \ldots, i_{n-1}$ (таких, что $p_{X_0, \ldots, X_{n-1}}(i_0, \ldots, i_{n-1}) > 0$) являются функциями только n, i_{n-1}, i_n. Отсюда легко следует, что (см. § 5)

$$p_{X_0, \ldots, X_n}(i_0, \ldots, i_n) = p_{X_0}(i) p_0(i_1 | i_0) \ldots p_{n-1}(i_n | i_{n-1}),$$

где

$$p_k(i_{k+1} | i_k) = p_{X_{k+1}|X_k}(i_{k+1} | i_k)$$

называются *переходными вероятностями марковской цепи в момент* k, $k=0, 1, 2, \ldots$. Если функция $p_k(j|i)$ зависит от k не формально, то марковская цепь называется *неоднородной*. С примерами неоднородных цепей мы встретились в § 11. Далее рассматриваем главным образом однородные марковские цепи, когда

$$p_k(j|i) = p(j|i) \equiv p_{ij}, \quad i, j \in \mathscr{I},$$

не зависят от k и называются просто *переходными вероятностями*.

Распределение вероятностей

$$p_{X_0}(i) = p_0(i) \equiv p_i^{(0)}, \quad i \in \mathscr{I},$$

называется *начальным распределением* марковской цепи; $p_0(i_0) = 1$, если процесс начинается из фиксированного состояния i_0.

Обозначим

$$p_i^{(n)} = p_{X_n}(i), \quad i \in \mathscr{I},$$

распределение вероятностей состояния системы в n-й момент времени, $n=0, 1, 2, \ldots$. Суммируя соотношение (см. 5.13)

$$p_{X_k, \ldots, X_{k+n}}(i_k, \ldots, i_{k+n}) = p_{X_k}(i_k) p(i_{k+1} | i_k) \ldots p(i_{k+n} | i_{k+n-1})$$

по всем возможным значениям $i_{k+1}, \ldots, i_{k+n-1}$, получим

$$p_{X_k, X_{k+n}}(i_k, i_{k+n}) = p_{X_k}(i_k) \sum_{\{i_{k+1}, \ldots, i_{k+n-1}\}} \prod_{m=k}^{k+n-1} p(i_{m+1} | i_m).$$

Условные вероятности

$$p_{X_{k+n}|X_k}(j|i) = p_{X_k, X_{k+n}}(i, j) / p_{X_k}(i) \equiv p^{(n)}(j|i) \equiv p_{ij}^{(n)}$$

называются *переходными вероятностями марковской цепи за n шагов;* в частности, $p_{ij}^{(1)} \equiv p_{ij}$. Из формулы

$$p_{ij}^{(n)} = \sum_{(i_1, i_2, \ldots, i_{n-1})} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{n-2} i_{n-1}} p_{i_{n-1} j}$$

легко следуют соотношения

$$p_{ik}^{(n+m)} = \sum_{(j)} p_{ij}^{(n)} p_{jk}^{(m)}, \quad i, \ k \in \mathcal{I},$$

называемые *уравнениями Колмогорова—Чепмена.* Если множество состояний $\mathcal{I}$ конечно, то, образовав матрицу $\mathcal{P}^{(n)} = \| p_{ij}^{(n)} \|$ из переходных вероятностей, получим

$$\mathcal{P}^{(n+m)} = \mathcal{P}^{(n)} \cdot \mathcal{P}^{(m)},$$

откуда видно, что матрица $\mathcal{P}^{(n)}$ переходных вероятностей за n шагов является n-й степенью матрицы $\mathcal{P} = \|p_{ij}\|$ переходных вероятностей за один шаг.

Суммируя основное соотношение

$$p_{X_0, X_1, \ldots, X_n}(i_0, \ i_1, \ \ldots, \ i_n) = p_{i_0}^{(0)} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}$$

по всем i_m, $0 \leqslant m \leqslant n$, кроме $i_{n_1}, \ldots, i_{n_k}$, $0 \leqslant n_1 < \ldots < n_k \leqslant n$, получим следующую формулу для конечномерных распределений процесса ($n(i) \equiv n_i$):

$$p_{X_{n(1)}, X_{n(2)}, \ldots, X_{n(k)}}(j_1, \ j_2, \ \ldots, \ j_k) = p_{j_1}^{(n_1)} p_{j_1 j_2}^{(n_2 - n_1)} \ldots p_{j_{k-1} j_k}^{(n_k - n_{k-1})}.$$

Наконец, отметим, что если p_{ij}, $i, \ j \in \mathcal{I}$, — произвольная совокупность чисел, таких, что $0 \leqslant p_{ij} \leqslant 1$ при всех $i, \ j$, и для любого $i \in \mathcal{I}$ $\{p_{ij}, \ j \in \mathcal{I}\}$ представляет собой распределение вероятностей на $\mathcal{I}$:

$$\sum_{(j \in \mathcal{I})} p_{ij} = 1,$$

то можно на некотором вероятностном пространстве $(\Omega, \mathcal{B}, \mathbf{P})$ определить марковскую цепь с переходными вероятностями p_{ij} и любым заданным начальным распределением (см. § 6).

14.1. Задача. Рассмотрим марковскую цепь с множеством всех целых чисел в качестве пространства состояний и переходными вероятностями p_{ij}, отличными от нуля только при $|i-j| = 1$. Положим $p_{i, i+1} = p_i$, $p_{i, i-1} = 1 - p_i$, $0 < p_i < 1$, при всех i. Пусть $a < b$ — некоторые целые числа, и обозначим $Q_{a,b}^{-}(k)$, $k \in (a, \ b)$, вероятность того, что марковская цепь, начиная свое движение из точки k, попадет раньше в a, чем в b. Вывести рекуррентное уравнение для $Q_{a,b}^{-}(k)$, $k \in (a, \ b)$, и решить его (ср. задачу о разорении игрока 2.20).

Обозначим $q_m(k)$ вероятность того, что траектория марковского блуждания попадет впервые в точку a на m-м шаге, не побывав до этого в точке b. Тогда

$$q_m(k)=\sum p_{ki_1}p_{i_1i_2}\dots p_{i_{m-1}a},$$

где суммирование ведется по всем *путям* $k\to i_1\to i_2\to\dots\to i_{m-1}\to a$ длины m, ведущим из точки k в точку a, минуя b. При $a+1<k<$ $<b-1$, $m>1$ запишем

$$q_m(k)=\sum(p_{k,k+1}p_{k+1,i_2}\dots p_{i_{m-1},a}+p_{k,k-1}p_{k-1,i_2}\dots p_{i_{m-1},a}).$$

Так как пути $k\pm1\to i_2\to\dots\to i_{m-1}\to a$ минуют b, то получаем отсюда

$$q_m(k)=p_kq_{m-1}(k+1)+(1-p_k)q_{m-1}(k-1).$$

При $k=a+1$ и $k=b-1$ аналогично получаем для $b-a>2$:

$$q_m(a+1)=p_{a+1}q_{m-1}(a+2),\quad m>1,\quad q_1(a+1)=1-p_{a+1},$$
$$q_m(b-1)=(1-p_{b-1})q_{m-1}(b-2),\quad m>1.$$

Суммируя полученные соотношения по m, получаем уравнение

$$\sum_{\{m\}}q_m(k)\equiv Q^-_{a,b}(k)=(1-p_k)Q^-_{a,b}(k-1)+p_kQ^-_{a,b}(k+1)$$

с граничными условиями $Q^-_{a,b}(a)=1$, $Q^-_{a,b}(b)=0$. Решая его (ср. 2.20) и полагая $\lambda_j=(1-p_j)/p_j$, получим при $a<k<b$:

$$Q^-_{a,b}(k)=\sum_{j=k}^{b-1}\lambda_{a+1}\lambda_{a+2}\dots\lambda_j/\left(1+\sum_{j=a+1}^{b-1}\lambda_{a+1}\lambda_{a+2}\dots\lambda_j\right)$$

Для вероятностей $Q^+_{a,b}(k)$ того, что траектория раньше попадет в точку b, чем в точку a, имеем то же уравнение, но с другими граничными условиями $Q^+_{a,b}(b)=1$, $Q^+_{a,b}(a)=0$, откуда получаем аналогично

$$Q^+_{a,b}(k)=\left(1+\sum_{j=a+1}^{k}\lambda_{a+1}\lambda_{a+2}\dots\lambda_j\right)\Big/\left(1+\sum_{j=a+1}^{b-1}\lambda_{a+1}\lambda_{a+2}\dots\lambda_j\right).$$

Отметим, что $Q^+_{a,b}(k)+Q^-_{a,b}(k)=1$, так что с вероятностью 1 траектория блуждания выходит на границу любого отрезка $[a, b]$ из любой внутренней точки k.]

14.2. Пример (неоднородное простое блуждание). Марковскую цепь задачи 14.1 называют иногда *неоднородным простым блужданием*. Следует, однако, иметь в виду, что если $p_k\neq p_{k'}$ хотя бы для одной пары k, k', то рассматриваемый процесс теряет одну важную черту простого блуждания — независимость приращений. Положение блуждающей частицы в момент n перестает быть представимым в виде суммы независимых единичных скачков (начиная с тех n, для которых достижимы из начального положения те области, где нарушена однородность вероятностей перехода). Вместе с тем, как мы видели в 14.1, техника исследования простого случайного блуждания отчасти работает и в неоднородном случае.

Разберем характер поведения траекторий неоднородного блуждания при $n \to \infty$. Пусть для определенности $X_0(\omega)=0$. Рассмотрим событие $C=\{\omega : \varliminf_{n\to\infty} X_n(\omega)=-\infty\}$. Обозначим $C_{a,b}$, $a<0<b$. (a, b — целые), событие, состоящее в том, что траектория $X_0(\omega)$, $X_1(\omega)$, $X_2(\omega), \ldots$ достигает точку a раньше, чем точку b. Последовательность $C_{a,b}$, $b=1, 2, \ldots$, возрастает к событию $C_a=\bigcup\limits_{b=1}^{\infty} C_{a,b}$, соответствующему всем таким траекториям, которые когда-то достигают положения a. По свойству непрерывности вероятности, используя 14.1 и принимая во внимание, что последовательность C_{-k}, $k=1, 2, \ldots$, убывает, $C=\bigcap\limits_{k=1}^{\infty} C_{-k}$, получаем

$$\mathbf{P}(C)=\lim_{a\to-\infty}\mathbf{P}(C_a),\ \mathbf{P}(C_a)=\lim_{b\to\infty}\mathbf{P}(C_{a,b})=$$

$$=\lambda_{a+1}\ldots\lambda_{-1}\cdot\Lambda^{+}\cdot(1+\lambda_{a+1}+\lambda_{a+1}\lambda_{a+2}+\ldots+\lambda_{a+1}\ldots\lambda_{-1}+$$

$$+\lambda_{a+1}\ldots\lambda_{-1}\Lambda^{+})^{-1},$$

$$\Lambda^{+}=\sum_{j=0}^{\infty}\lambda_0\lambda_1\ldots\lambda_j.$$

Отсюда следует, что при $\Lambda^+<\infty$ $\mathbf{P}(C_a)<1$, $\mathbf{P}(C)<1$, а при $\Lambda^+=\infty$, $\mathbf{P}(C_a)=\mathbf{P}(C)=1$.

Если начальная точка блуждания отлична от нуля, то вопрос сводится к сходимости или расходимости ряда

$$\sum_{j=k}^{\infty}\lambda_k\lambda_{k+1}\ldots\lambda_j,\quad (X_0(\omega)=k),$$

который имеет вид $\alpha_k+\beta_k\Lambda^+$ при некоторых постоянных α_k, β_k. Это означает, что $\mathbf{P}(C)=1$ или $\mathbf{P}(C)<1$ одновременно для всех начальных положений (или для любого начального распределения).

Характерно, что возможность побывать в сколь угодно далеких *отрицательных* точках определяется поведением вероятностей перехода p_k, $1-p_k$ в сколь угодно далеких *положительных* точках числовой прямой. Расходимость ряда Λ^+ означает, выражаясь приблизительно, что при больших положительных j $\lambda_j=(1-p_j)/p_j$ в своем большинстве $\geqslant 1$, т. е. блуждающая точка имеет тенденцию из области больших положительных значений возвращаться в какую-то конечную окрестность начала координат. В частном случае $p_k=p=\text{const}$ при $k>0$ (или при всех k начиная с некоторого) $\Lambda^+=\infty$ тогда и только тогда, когда $(1-p)/p\geqslant 1$, т. е. $p\leqslant 1/2$.

Ввиду симметрии задачи получаем, что событие $\{\omega : \overline{\lim\limits_{n\to\infty}} X_n(\omega)=\infty\}$ имеет вероятность 1 или <1 в зависимости от расходимости или сходимости ряда

$$\Lambda^-=\sum_{j=0}^{\infty}(\lambda_0\lambda_{-1}\ldots\lambda_{-j})^{-1},\quad \lambda_k=(1-p_k)/p_k.$$

Если $\Lambda^-=\Lambda^+=\infty$, то блуждание является *осциллирующим:*

$$\mathbf{P}(\omega : \underline{\lim_{n\to\infty}} X_n(\omega)=-\infty,\ \overline{\lim_{n\to\infty}} X_n(\omega)=\infty)=1.$$

Докажем, что если $\Lambda^+=\infty$, $\Lambda^-<\infty$, то с вероятностью 1 $X_n(\omega)\to-\infty$ при $n\to\infty$. Пусть для определенности $X_0(\omega)=1$. Обозначим через $T_k(\omega)$ — момент k-го возвращения траектории X_0, X_1, $X_2,\ldots$ в точку 0 для таких ω, для которых такой момент существует, и положим $T_k(\omega)=\infty$ для остальных ω (т. е. для тех ω, для которых последовательность $X_1(\omega)$, $X_2(\omega),\ldots$ содержит $<k$ значений, равных 0).

Найдем вероятность события $\{\omega : T_1(\omega)<\infty\}$. Имеем

$$\mathbf{P}(\omega : T_1(\omega)=n)=\sum p_{0i_1}p_{i_1i_2}\ldots p_{i_{n-1}0},$$

где суммирование ведется по всем значениям i_1, $i_2,\ldots,i_{n-1}$, не равным нулю. Разобьем сумму на две

$$p_{01}\sum p_{1i_2}p_{i_2i_3}\ldots p_{i_{n-1}0}+p_{0,-1}\sum p_{-1,i_2}p_{i_2i_3}\ldots p_{i_{n-1}0}$$

и заметим, что $p_{0,1}=p_0$, $p_{0,-1}=1-p_0$, а

$$\sum p_{1i_1}\ldots p_{i_{n-1}0}\equiv q'_{n-1}(1),\quad \sum p_{-1,i_1}\ldots p_{i_{n-1}0}\equiv q''_{n-1}(-1)$$

представляют собой соответственно вероятность того, что блуждание, выходящее из точки 1, попадает впервые в точку 0 на $(n-1)$-м шаге, и вероятность такого же события, но для блуждания, начинающегося из точки -1. Из условия $\Lambda^+=\infty$ вытекает, что

$$\sum_{n=2}^{\infty} q'_{n-1}(1)=1,$$

а из условия $\Lambda^-<\infty$ следует, что

$$\sum_{n=2}^{\infty} q''_{n-1}(-1)<1,$$

откуда

$$\mathbf{P}(\omega : T_1(\omega)<\infty)=p_0+(1-p_0)\sum_{n=2}^{\infty} q''_{n-1}(-1)\equiv\theta<1.$$

Покажем индукцией по n, что

$$\mathbf{P}(\omega: T_n(\omega)<\infty)=\theta^n.$$

Предположим, что для $n-1$ утверждение установлено. Тогда, действуя подобно предыдущему, запишем

$$\mathbf{P}(\omega: T_n(\omega)<\infty)=\sum p_{0i_1}\dots p_{i_{k-1}0}p_{0i_{k+1}}\dots p_{i_{k+m}0},$$

где суммирование ведется по всевозможным k, m, $i_1,\dots,i_{k+m}$, таким, что $i_1,\dots,i_{k-1}$ отличны от нуля, а среди чисел $i_{k+1},\dots,i_{k+m}$ ровно $n-2$ нуля. Разбивая сумму на повторную, получаем

$$\sum_{\{k,i_1,\dots,i_{k-1}\}} p_{0i_1}\dots p_{i_{k-1}0} \sum_{\{m,\ i_{k+1},\dots,i_{k+m}\}} p_{0i_{k+1}}\dots p_{i_{k+m}0} =$$

$$=\mathbf{P}(\omega: T_1(\omega)<\infty)\cdot\mathbf{P}(\omega: T_{n-1}(\omega)<\infty)=\theta^n$$

Обозначив $N_0(\omega)\leqslant\infty$ число всех возвращений блуждания в начальную точку, имеем

$$\mathbf{P}(\omega: N_0(\omega)\geqslant n)=\mathbf{P}(\omega: T_n(\omega)<\infty)=\theta^n,$$

так что $\mathbf{P}(\omega: N_0(\omega)<\infty)=1$. Так же и для любой другой точки k число $N_k(\omega)$ попаданий блуждания в точку k конечно с вероятностью 1. Проверим это для $k<0$. Обозначим $\tau_k(\omega)$ момент первого попадания в точку k. Поскольку $\lim\limits_{n\to\infty} X_n(\omega)=-\infty$ с вероятностью 1, то $\tau_k(\omega)<\infty$ для любого $k<0$ с вероятностью 1. Легко видеть, что последовательность сл. в. $Y_n=X_{\tau_k+n}$, $n=0, 1, 2, \dots$, представляет собой неоднородное простое блуждание, начинающееся из точки k и не зависящее от отрезка блуждания $X_0, X_1, \dots, X_{\tau_k}$. В самом деле, при $k_0=0$, $l_0\equiv k$ имеем

$$\mathbf{P}(\omega: X_i(\omega)=k_i,\ i<m,\ \tau_k(\omega)=m,\ Y_j(\omega)=l_j,\ j\leqslant n)=$$

$$=p_{0k_1}p_{k_1k_2}\dots p_{k_{m-1}k}p_{kl_1}p_{l_1l_2}\dots p_{l_{n-1}l_n},\quad k_i\neq k,\ i<m.$$

Суммируя это соотношение по всем возможным m, $k_1,\dots,k_{m-1}$ (где последовательность $k_1,\dots,k_{m-1}$ не содержит точки k), получим

$$\mathbf{P}(\omega: Y_j(\omega)=l_j,\ j\leqslant n)=p_{kl_1}p_{l_1l_2}\dots p_{l_{n-1}l_n}\quad (l_0=k),$$

а суммируя по $l_1,\dots,l_n$, находим

$$\mathbf{P}(\omega: X_i(\omega)=k_i,\ i<m,\ \tau_k(\omega)=m)=p_{0k_1}p_{k_1k_2}\dots p_{k_{m-1}k},$$

так что сл. в. Y_n, $n=0, 1, 2,\dots$, не зависят от τ_k, $X_0, X_1, \dots, X_{\tau_k}$. Поскольку сл. в. $N_k(\omega)$ равняется сл. в. $N_0'(\omega)+1$, где $N_0'(\omega)$ — число возвращений в начальное состояние марковской цепи $Y_n(\omega)$, $n=0, 1, 2,\dots$, то $\mathbf{P}(\omega: N_k(\omega)<\infty)=\mathbf{P}(\omega: N_0'(\omega)<\infty)=1$.

Таким образом, мы получили, что с вероятностью 1 $\underline{\lim}_{n\to\infty} X_n \times$ $\times (\omega) = -\infty$, $N_k(\omega) < \infty$, $k = -1, -2, \ldots$. Положим

$$X(\omega) = \overline{\lim}_{n\to\infty} X_n(\omega).$$

Допустим, что при каком-либо $k<0$ $\mathbf{P}(\omega : X(\omega) > k) > 0$. Тогда с положительной вероятностью $N_k(\omega) = \infty$. Полученное противоречие показывает, что $X(\omega) = -\infty$ с вероятностью 1, что и требовалось доказать.

Ввиду симметрии задачи получаем, что при $\Lambda^+ < \infty$, $\Lambda^- = \infty$

$$\mathbf{P}(\omega : \lim_{n\to\infty} X_n(\omega) = \infty) = 1.$$

Покажем, что если $\Lambda^+ < \infty$, $\Lambda^- < \infty$, то для п. в. траекторий имеет место одно из двух

$$\lim_{n\to\infty} X_n(\omega) = \infty \text{ или } \lim_{n\to\infty} X_n(\omega) \leqslant -\infty,$$

причем обе возможности имеют положительные вероятности. Заметим, что так же, как и выше, устанавливается конечность с вероятностью 1 сл. в. $N_k(\omega)$ — числа попаданий траектории в точку k, $k=0, \pm 1, \pm 2, \ldots$. Следовательно, на любом отрезке $[a, b]$ траектория пребывает с вероятностью 1 конечное время, а потому, во-первых, каждое из событий

$$\underline{\lim}_{n\to\infty} X_n(\omega) > -\infty, \quad \overline{\lim}_{n\to\infty} X_n(\omega) < \infty$$

имеет нулевую вероятность и, во-вторых, события

$$\underline{\lim}_{n\to\infty} X_n(\omega) = -\infty, \quad \overline{\lim}_{n\to\infty} X_n(\omega) = \infty$$

одновременно выполняются лишь с нулевой вероятностью. Таким образом, для п. в. траекторий выполняется ровно одна из двух возможностей

$$\lim_{n\to\infty} X_n(\omega) = \infty \text{ или } \lim_{n\to\infty} X_n(\omega) = -\infty.$$

Каждая из этих возможностей осуществляется с положительной вероятностью, так как ввиду условий $\Lambda^+ < \infty$, $\Lambda^- < \infty$

$$\mathbf{P}(\omega : X_i(\omega) < 1,\ i=0,\ 1,\ 2,\ \ldots) > 0,$$

$$\mathbf{P}(\omega : X_i(\omega) > -1,\ i=0,\ 1,\ 2,\ \ldots) > 0.$$

Отметим, что при $p_k = p = \text{const}$ при всех $k=0, \pm 1, \pm 2, \ldots$ ряды Λ^+ и Λ^- представляют собой суммы геометрических прогрессий со знаменателем $(1-p)/p$ и $p/(1-p)$ соответственно. Отсюда получаем ранее уже установленные результаты о поведении простого (однородного) случайного блуждания.

14.3. Задача. Рассмотрим простое неоднородное блуждание, начинающееся из какой-либо точки $k \in [a, b]$, a, b — целые, и обозначим $\tau_k(\omega)$ момент первого выхода блуждания из отрезка $[a, b]$. Положим $r=[(b-a)/2]+1$ и

$$\theta=\max_{k\in[a,b]} \mathbf{P}(\omega:\tau_k(\omega)>r).$$

Показать, что

$$\mathbf{P}(\omega:\tau_k(\omega)>jr)\leqslant\theta^j,\quad j=1,\ 2,\ \ldots,$$

разбивая траекторию блуждания на участки длиной r.

Отметим прежде всего, что $\theta<1$. В самом деле, для блуждания, начинающегося из точки $k\in[a, b]$, один из путей $k\to k+1\to\ldots\to k+r$ и $k\to k-1\to\ldots\to k-r$ выводит за пределы отрезка $[a, b]$. Обозначив $\alpha_k>0$ вероятность такого пути, имеем $\min\limits_{k\in[a,b]}\alpha_k\equiv\alpha>0$, и для любого $k\in[a, b]$

$$\mathbf{P}(\omega:\tau_k(\omega)>r)=1-\mathbf{P}(\omega:\tau_k(\omega)\leqslant r)\leqslant 1-\alpha.$$

Далее запишем при $j>1$

$$\mathbf{P}(\omega:\tau_k(\omega)>jr)=\sum p_{ki_1}p_{i_1i_2}\ldots p_{i_{jr-1}i_{jr}},$$

где суммирование ведется по всем путям $k\to i_1\to i_2\to\ldots\to i_{jr}$, не выходящим за пределы отрезка $[a, b]$. Разбивая все такие пути на два участка

$$k\to i_1\to i_2\to\ldots\to i_{(j-1)r},\ i_{(j-1)r}\to i_{(j-1)r+1}\to\ldots\to i_{jr},$$

запишем сумму как повторную $(p(j|i)\equiv p_{ij})$:

$$\sum p(i_1|k)\ldots p(i_{(j-1)r}|i_{(j-1)r-1})\sum p(i_{(j-1)r+1}|i_{(j-1)r})\ldots p(i_{jr}|i_{jr-1}),$$

где внешнее суммирование ведется по путям из первого участка, а внутреннее — по путям из второго. Поскольку внутренняя сумма равна вероятности

$$\mathbf{P}(\omega:\tau_{i_{(j-1)r}}>r)\leqslant\theta,$$

то, заменяя ее на θ и вынося оценку за знак внешней суммы, получим, что остающаяся сумма есть

$$\mathbf{P}(\omega:\tau_k(\omega)>(j-1)r),$$

так что

$$\mathbf{P}(\omega:\tau_k(\omega)>jr)\leqslant\theta\cdot\mathbf{P}(\omega:\tau_k(\omega)>(j-1)r).$$

Применяя индукцию, получаем требуемый результат.]

14.4. Задача. Обозначим m_k среднее время выхода неоднородного простого блуждания, начинающегося в точке $k\in[a, b]$, из отрезка $[a, b]$. Вывести рекуррентное уравнение для m_k.

[Отметим, что ввиду 14.3 для хвостов распределения времени τ_k выхода из отрезка $[a, b]$ имеем оценку

$$\mathbf{P}(\omega:\tau_k(\omega)>t)\leqslant\mathbf{P}(\omega:\tau_k(\omega)>[t/r]r)\leqslant\theta^{[t/r]},\quad r=[(b-a)/2]+1,$$

так что $m_k<\infty$. Далее запишем при целом t

$$\mathbf{P}(\omega:\tau_k(\omega)>t)=\sum p_{ki_1}p_{i_1i_2}\dots p_{i_{t-1}i_t},$$

где суммирование ведется по всем целым $i_1,\dots,i_t$ из отрезка $[a, b]$. Разбивая при $a+1<k<b-1$ сумму на две, приведем ее к виду

$$p_k\sum p_{k+1,i_2}\dots p_{i_{t-1}i_t}+(1-p_k)\sum p_{k-1,i_2}\dots p_{i_{t-1}i_t}=$$
$$=p_k\mathbf{P}(\omega:\tau_{k+1}(\omega)>t-1)+(1-p_k)\mathbf{P}(\omega:\tau_{k-1}(\omega)>t-1).$$

Суммируя по t, получим (см. 7.3)

$$m_k=\sum_{t=0}^{\infty}\mathbf{P}(\omega:\tau_k(\omega)>t)=1+\sum_{t=1}^{\infty}\mathbf{P}(\omega:\tau_k(\omega)>t)=$$

$$=1+p_k\sum_{t=1}^{\infty}\mathbf{P}(\omega:\tau_{k+1}(\omega)>t-1)+(1-p_k)\sum_{t=1}^{\infty}\mathbf{P}(\omega:\tau_{k-1}(\omega)>t-1),$$

или

$$m_k=1+p_km_{k+1}+(1-p_k)m_{k-1},\quad a<k<b.$$

Точно так же выводятся соотношения для $k=a$ и $k=b$, которые сводятся к предыдущему, если положить $m_{a-1}=m_{b+1}=0$.]

Говорят, что m_k, $a\leqslant k\leqslant b$, является решением уравнения $m_k=1+p_km_{k+1}+(1-p_k)m_{k-1}$ с граничными условиями $m_{a-1}=m_{b+1}=0$. Решение этого уравнения проводится в задачах 14.5—14.7 и основано на следующей общей идее. Пусть m_k^* — какое-то решение указанного уравнения (которое было бы легко найти в явной форме), m_k — искомое решение и $x_k=m_k-m_k^*$. Очевидно, x_k удовлетворяет однородному уравнению

$$x_k=p_kx_{k+1}+(1-p_k)x_{k-1}$$

с граничными условиями $x_{a-1}=-m_{a-1}^*$, $x_{b+1}=-m_{b+1}^*$, где m_{a-1}^*, m_{b+1}^* нам известны. Но решение этого уравнения с заданными граничными условиями уже не составляет труда (см. 14.1 и 2.20).

14.5. Задача. Найти решение m_k^* уравнения задачи 14.4 с граничными условиями $m_{a-1}=m_a=0$.

[Переписывая уравнение в виде

$$m_k-m_{k+1}=\lambda_k(m_{k-1}-m_k)+p_k^{-1},\quad \lambda_k=(1-p_k)p_k^{-1},$$

и учитывая условия $m_{a-1}=m_a=0$, получаем для его решения m_k^* соотношения

$$m_k^*-m_{k+1}^*=\sum_{j=a}^{k-1}\lambda_k\lambda_{k-1}\dots\lambda_{j+1}p_j^{-1}+p_k^{-1},\quad k\geqslant a,$$

где сумму по пустому множеству индексов считаем равной нулю. Суммируя по k от a до $i-1$, находим

$$-m_i^* = \sum_{k=a}^{i-1} \left(\sum_{j=a}^{k-1} \lambda_k \lambda_{k-1} \ldots \lambda_{j+1} p_j^{-1} + p_k^{-1} \right), \quad i \geqslant a. \;]$$

14.6. Задача. Найти решение уравнения

$$x_k = p_k x_{k+1} + (1-p_k) x_{k-1}, \quad k \in [a,\ b],$$

с граничными условиями $x_{a-1}=0$, $x_{b+1}=x$.

[Аналогично 14.5 имеем

$$x_k - x_{k+1} = \lambda_k \lambda_{k-1} \ldots \lambda_a x_a, \quad k \geqslant a,$$

$$x_i = \left(1 + \sum_{k=a}^{i-1} \lambda_k \lambda_{k-1} \ldots \lambda_a \right) x_a, \quad i \geqslant a,$$

$$x = x_{b+1} = \left(1 + \sum_{k=a}^{b} \lambda_k \lambda_{k-1} \ldots \lambda_a \right) x_a,$$

$$x_i = \left(1 + \sum_{k=a}^{i-1} \lambda_k \lambda_{k-1} \ldots \lambda_a \right) \left(1 + \sum_{k=a}^{b} \lambda_k \lambda_{k-1} \ldots \lambda_a \right)^{-1} x. \;]$$

14.7. Задача. Найти решение m_k уравнения задачи 14.4.

Искомое решение равно $m_k = m_k^* + x_k$, где m_k^* определено в 14.5, а x_k — в 14.6 при $x = -m_{b+1}^*$, так что

$$m_k = -\sum_{i=a}^{k-1} \left(\sum_{j=a}^{i-1} \lambda_i \lambda_{i-1} \ldots \lambda_{j+1} p_j^{-1} + p_i^{-1} \right) + \left(1 + \sum_{i=a}^{k-1} \lambda_i \lambda_{i-1} \ldots \lambda_a \right) \times$$

$$\times \left(1 + \sum_{i=a}^{b} \lambda_i \lambda_{i-1} \ldots \lambda_a \right)^{-1} \sum_{i=a}^{b} \left(\sum_{j=a}^{i-1} \lambda_i \lambda_{i-1} \ldots \lambda_{j+1} p_j^{-1} + p_i^{-1} \right), \quad k \geqslant a,$$

с тем соглашением, что сумма по пустому множеству индексов равна нулю. Для последующего использования перепишем эту формулу более компактно, полагая $\sigma_s = \lambda_s \lambda_{s-1} \ldots \lambda_a$, $s \geqslant a$, $\sigma_{a-1} = 1$:

$$m_k = -\sum_{i=a}^{k-1} \sigma_i \sum_{j=a}^{i} \sigma_j^{-1} p_j^{-1} + \sum_{i=a-1}^{k-1} \sigma_i \left(\sum_{i=a-1}^{b} \sigma_i \right)^{-1} \sum_{i=a}^{b} \sigma_i \sum_{j=a}^{i} \sigma_j^{-1} p_j^{-1}, \quad k \geqslant a. \;]$$

14.8. Задача. Рассмотрим задачу о разорении игрока 2.20, полагая в предыдущих соотношениях $p_k = p = \text{const}$, $a=1$, заменяя b на $b-1$, так что окончание игры происходит, когда траектория простого однородного блуждания, выходящая из точки $k \in [1,\ b-1]$, попадает в одну из точек 0 или b. Вывести из 14.7 следующие формулы для среднего времени продолжительности игры

$$m_k=k(b-k),\quad p=\frac{1}{2};\quad m_k=\frac{k}{1-2p}-$$

$$-\frac{b}{1-2p}\cdot\frac{1-((1-p)/p)^k}{1-((1-p)/p)^b},\quad p\neq\frac{1}{2}.$$

$$\Bigg[m_k=-\sum_{i=1}^{k-1}\sum_{j=1}^{i}\left(\frac{1-p}{p}\right)^{i-j}\frac{1}{p}+\sum_{i=0}^{k-1}\left(\frac{1-p}{p}\right)^{i}\times$$

$$\times\left(\sum_{i=0}^{b-1}\left(\frac{1-p}{p}\right)^{i}\right)^{-1}\cdot\sum_{i=1}^{b-1}\sum_{j=1}^{i}\left(\frac{1-p}{p}\right)^{i-j}\frac{1}{p}.$$

Пусть, например, $p\neq 1/2$, тогда первое и второе слагаемые в выражении для m_k преобразуются соответственно к виду

$$-\frac{1}{1-2p}\sum_{i=1}^{k-1}\left(\left(\frac{1-p}{p}\right)^{i}-1\right)=\frac{k}{1-2p}-$$

$$-\left(\left(\frac{1-p}{p}\right)^{k}-1\right)\frac{p}{(1-2p)^2},$$

$$\left(\left(\frac{1-p}{p}\right)^{k}-1\right)\left(\left(\frac{1-p}{p}\right)^{b}-1\right)^{-1}\times$$

$$\times\left(-\frac{b}{1-2p}+\left(\left(\frac{1-p}{p}\right)^{b}-1\right)\frac{p}{(1-2p)^2}\right).$$

Складывая, получаем требуемое выражение.]

14.9. Пример (среднее время достижения в неоднородном блуждании). Рассмотрим случай $\Lambda^+=\sum_{i=0}^{\infty}\lambda_0\lambda_1\ldots\lambda_i=\infty$, когда с вероятностью 1 блуждание достигает любой точки слева от начальной. Введем сл. в. $T_k^{a,b}$, T_k^a, равные соответственно времени до выхода из отрезка $[a, b]$ и первого попадания в точку $a-1$ для блуждания, начинающегося из точки $k\in[a, b]$. Последовательность сл. в. $T_k^{a,b}$, $b=k+1, k+2, \ldots$, не убывает и с вероятностью 1 сходится к T_k^a. Теория, изложенная в § 10, устанавливает законность предельного перехода под знаком математического ожидания

$$\lim_{b\to\infty}\mathbf{M}T_k^{a,b}=\mathbf{M}T_k^a\leqslant\infty.$$

Следовательно, для вычисления $\mathbf{M}T_k^a$ можно перейти к пределу в формуле 14.7 для $\mathbf{M}T_k^{a,b}=m_k$. Предположим, что сходится ряд

$$\sum_{j=a}^{\infty}\sigma_j^{-1}p_j^{-1}=S<\infty.$$

Обозначим S_n его частичную сумму и положим

$$A_n=\sum_{i=a}^{n}\sigma_i S_i \Big/ \sum_{i=a-1}^{n}\sigma_i,$$

так что

$$MT_k^a=-\sum_{i=a}^{k-1}\sigma_i\sum_{j=a}^{i}\sigma_j^{-1}p_j^{-1}+\sum_{i=a-1}^{k-1}\sigma_i\lim_{b\to\infty}A_b$$

(где сумма по пустому множеству индексов полагается равной нулю). Поскольку $\Lambda^+=\infty$, то

$$\lim_{n\to\infty}\sum_{i=0}^{n}\sigma_i=\infty,$$

а ввиду конечности предела частичных сумм S_n легко вывести, что

$$\lim_{n\to\infty}A_n=\lim_{n\to\infty}S_n=S.$$

Таким образом, с учетом $\sigma_{a-1}=1$, $\sigma_i=\lambda_i\lambda_{i-1}\ldots\lambda_a$, $i\geqslant a$, имеем

$$MT_k^a=-\sum_{i=a}^{k-1}\sigma_i\sum_{j=a}^{i}\sigma_j^{-1}p_j^{-1}+\sum_{i=a-1}^{k-1}\sigma_i\sum_{j=a}^{\infty}\sigma_j^{-1}p_j^{-1}=\sum_{i=a}^{k-1}\sigma_i\sum_{j=i+1}^{\infty}\sigma_j^{-1}p_j^{-1}+$$

$$+\sum_{j=a}^{\infty}\sigma_j^{-1}p_j^{-1}=\sum_{i=a-1}^{k-1}\sigma_i\sum_{j=i+1}^{\infty}\sigma_j^{-1}p_j^{-1}=\sum_{i=a-1}^{k-1}\sum_{j=i+1}^{\infty}(\lambda_{i+1}\lambda_{i+2}\ldots\lambda_j)^{-1}p_j^{-1}$$

Из полученной формулы, в частности, видно, что среднее время достижения точки $a-1$ из точки k складывается из суммы средних времен достижения точки i из точки $i+1$ по i от $k-1$ до $a-1$. Это, впрочем, легко понять и без вычислений. Пусть блуждание выходит из точки k, обозначим $\tau_j(\omega)$ момент первого достижения точки $k-j$, $j=1, 2, \ldots$. Стандартные рассуждения показывают. что при любом j отрезок блуждания $X_0, X_1, \ldots, X_{\tau_j}$ не зависит от последовательности сл. в. $Y_n=X_{\tau_j+n}$, $n=0, 1, 2, \ldots$, которая представляет собой блуждание, начинающееся из точки $k-j$ (см. 14.2). Сл. в. $\tau_{j+1}-\tau_j$ есть время до первого достижения блужданием Y_n, $n=0, 1, 2, \ldots$, точки $k-j-1$, а сл. в.

$$T_k^a\equiv\tau_{k-a+1}=\sum_{j=0}^{k-a}(\tau_{j-1}-\tau_j),\quad \tau_0=0,$$

представима в виде суммы случайных величин, имеющих распределения вероятностей времен достижения состояния $i-1$ из i по i от k до a. Применяя операцию математического ожидания, получаем требуемое.

Если условие $S<\infty$ не выполняется, то $\lim\limits_{n\to\infty}A_n=\infty$ и $MT_k^a=\infty$

при всех $k \geqslant a$. В частном случае однородного блуждания $p_k = p < 1/2$ имеем для среднего времени достижения соседней точки слева

$$MT_a^a = \sum_{j=a}^{\infty} \sigma_j^{-1} p_j^{-1} = 1/(1-2p),$$

а при $p = 1/2$, $MT_a^a = \infty$. Эти результаты были получены ранее в § 11, 12. Интересно посмотреть, что дает в неоднородном случае метод, примененный для нахождения MT_a^a в однородном блуждании. Именно, пусть $\mu_k \leqslant \infty$ есть среднее время достижения точки $k-1$ из точки k. Если первый шаг сделан налево, то на всех таких траекториях время достижения точки $k-1$ равно 1, а суммарная вероятность этих траекторий равна $1-p_k$, так что вклад в выражение для μ_k составляет $1 \cdot (1-p_k)$. Если первый шаг сделан направо, то, чтобы достигнуть точки $k-1$, надо сначала из $k+1$ попасть в k, а затем из k в $k-1$. Среднее время равно сумме средних $\mu_{k+1} + \mu_k$, а учитывая и первый шаг направо, получаем уравнение

$$\mu_k = 1 - p_k + p_k(1 + \mu_k + \mu_{k+1}),$$

или

$$\mu_k = 1/(1-p_k) + (p_k/(1-p_k))\mu_{k+1}.$$

Итерируя, находим

$$\mu_k = \sum_{j=k}^{\infty} \sigma_j^{-1} p_j^{-1} + \sigma_n^{-1} \mu_{n+1}.$$

Эти рассуждения легко сделать вполне аккуратными, используя уже применявшуюся выше технику вывода рекуррентных соотношений, однако, чтобы получить отсюда верную формулу

$$\mu_k = \sum_{j=k}^{\infty} \sigma_j^{-1} p_j^{-1} \equiv MT_k^k,$$

требуется отдельно доказывать, что $\sigma_n^{-1}\mu_{n+1} \to 0$ в предположении, что ряд для μ_k сходится.

14.10. Пример (процессы Гальтона—Ватсона). Пусть p_j, $j = 0, 1, 2, \ldots,$ — распределение вероятностей на множестве целых неотрицательных чисел, $\mathscr{P}(s) = \sum_{j=0}^{\infty} s^j p_j$, $0 \leqslant s \leqslant 1$, — его производящая функция (п. ф.). Определим марковскую цепь с множеством состояний $\mathscr{S} = \{0, 1, 2, \ldots\}$ и переходными вероятностями, определяемыми из соотношений

$$\sum_{j=0}^{\infty} s^j p_{ij} = (\mathscr{P}(s))^i, \quad i = 1, 2, \ldots, \quad p_{00} = 1.$$

При $i>0$ распределение вероятностей p_{ij}, $j=0, 1, 2, \ldots$, представляет собой распределение суммы i независимых одинаково распределенных случайных величин с п. ф. $\mathcal{P}(s)$. Марковскую цепь $X_0(\omega)$, $X_1(\omega), \ldots$ с переходными вероятностями p_{ij}, определенными выше, интерпретируют как некоторый процесс размножения частиц, при котором $X_{n+1}(\omega)$ есть число потомков $X_n(\omega)$ частиц предыдущего поколения, которые размножаются независимо и с одним и тем же распределением вероятностей p_j, $j=0, 1, 2, \ldots$.

Найдем п. ф. вероятностей перехода за n шагов. При $n=2$ имеем

$$\sum_{j=0}^{\infty} p_{ij}^{(2)}s^j=\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_{ik}p_{kj}s^j=\sum_{k=0}^{\infty} p_{ik}\sum_{j=0}^{\infty} p_{kj}s^j=\sum_{k=0}^{\infty} p_{ik}\mathcal{P}(s)^k=(\mathcal{P}(\mathcal{P}(s)))^i.$$

По индукции получаем общую формулу

$$\sum_{j=0}^{\infty} p_{ij}^{(n)}s^j=(\mathcal{P}(\mathcal{P}(\ldots\mathcal{P}(s)\ldots)))^i, \quad i\geqslant 1,$$

где справа стоит i-кратная суперпозиция функции $\mathcal{P}(s)$.

Состояние 0 рассматриваемого *ветвящегося процесса Гальтона — Ватсона* играет особую роль. Марковская цепь, раз попав в это состояние, остается там навсегда. Такое состояние называют *поглощающим*. Когда ветвящийся процесс попадает в поглощающее состояние, говорят, что он *выродился*. Вероятность, что процесс, начавшийся с i частиц, выродился до момента n включительно, равна

$$p_{i0}^{(n)}=(\mathcal{P}(\mathcal{P}(\ldots(\mathcal{P}(0))\ldots)))^i=(p_{10}^{(n)})^i.$$

Последовательность событий $A_n=\{\omega: X_n(\omega)=0\}$ возрастает к событию $A=\bigcup_{n=1}^{\infty} A_n$, означающему, что процесс когда-либо вырождается. По свойству непрерывности вероятности получаем

$$\mathbf{P}(A)=\lim_{n\to\infty} p_{i0}^{(n)}, \quad X_0(\omega)=i.$$

Чтобы избежать тривиальностей, предполагаем далее, что $0<$ $<p_0$, $p_0+p_1<1$, т. е. с положительной вероятностью частица погибает, не дав потомства, и есть положительная вероятность иметь двух или более потомков. Производящая функция $\mathcal{P}(s)$ — выпуклая вниз функция, и, как видно из рис. 35, последовательные итерации $\mathcal{P}_n(s)=\mathcal{P}(\mathcal{P}(\ldots\mathcal{P}(s)\ldots))$ сходятся при $n\to\infty$ и $0\leqslant s<1$ к точке s_0 — наименьшему корню уравнения $\mathcal{P}(s)=s$ на отрезке $[0, s]$. Проверим это для $s=0$.

Во-первых, так как производящая функция — возрастающая, то $\mathcal{P}_{n+1}(0)=\mathcal{P}_n(\mathcal{P}(0))>\mathcal{P}_n(0)$ и, следовательно, существует $\lim\limits_{n\to\infty}\mathcal{P}_n(0)=p$. Переходя к пределу в равенстве $\mathcal{P}_{n+1}(0)=\mathcal{P}(\mathcal{P}_n(0))$, получаем, что $\mathcal{P}(p)=p$. Если $s_0=1$, то равенство $p=s_0$ очевидно. Если же $s_0<1$, то, как нетрудно проверить, $\mathcal{P}'(s_0)<1$ и при $s<s_0$

$$0<s_0-\mathcal{P}(s)=\mathcal{P}(s_0)-\mathcal{P}(s)\leqslant\mathcal{P}'(s_0)(s_0-s).$$

Итерируя это соотношение, получим

$$0<s_0-\mathcal{P}_n(s)\leqslant\mathcal{P}'(s_0)^n(s_0-s).$$

При $s=0$ и $n\to\infty$ выводим $s_0=p$.

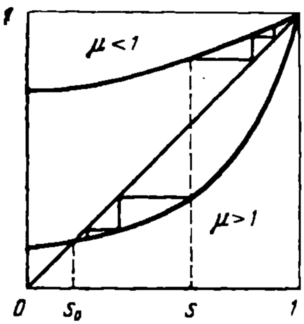

Рис. 35

Если $\mathcal{P}'(1-)>1$, то *вероятность вырождения* $p<1$. Если $\mathcal{P}'(1-)\leqslant1$, то $\mathcal{P}'(s)<1$ при $s<1$. Интегрируя это неравенство, получаем, что $\mathcal{P}(s)>s$, $s<1$, и, следовательно, вероятность вырождения $p=1$. Число $\mu=\mathcal{P}'(1-)$ есть математическое ожидание числа потомков одной частицы в первом поколении. Таким образом, при $\mu\leqslant1$ процесс вырождается с вероятностью 1 при любом начальном числе частиц (или произвольном начальном распределении), при $\mu>1$ с положительной вероятностью процесс никогда не попадает в поглощающее состояние.

14.11. Задача. Пусть $X_0(\omega)=1$, $X_1(\omega),\ldots,$ — ветвящийся процесс с п. ф. $\mathcal{P}(s)$. Предполагая, что $\mathcal{P}'(1-)$ и $\mathcal{P}''(1-)$ конечны, вычислить $\mathbf{M}X_n$, $\mathbf{D}X_n$.

[Дифференцируя соотношение $\mathcal{P}_{n+1}(s)=\mathcal{P}(\mathcal{P}_n(s))$, имеем

$$\mathcal{P}'_{n+1}(s)=\mathcal{P}(\mathcal{P}_n(s))'=\mathcal{P}'(\mathcal{P}_n(s))\mathcal{P}'_n(s),$$

$$\mathcal{P}''_{n+1}(s)=\mathcal{P}''(\mathcal{P}_n(s))\mathcal{P}'_n(s)^2+\mathcal{P}'(\mathcal{P}_n(s))\mathcal{P}''_n(s)).$$

Переходя к пределу при $s\to1$, $s<1$, находим

$$\mathcal{P}'_{n+1}(1-)=\mathcal{P}'(1-)\mathcal{P}'_n(1-),\quad \mathcal{P}'_n(1-)=\mathcal{P}'(1-)^n=\mu^n,$$

$$\mathcal{P}''_{n+1}(1-)=\mathcal{P}''(1-)\mathcal{P}'_n(1-)^2+\mathcal{P}'(1-)\mathcal{P}''_n(1-),$$

$$\mathcal{P}''_n(1-)=\mathcal{P}''(1-)\sum_{k=1}^{n}\mu^{k-1}\mu^{2(n-k)},$$

откуда (см. 8.6)

$$\mathcal{P}''_n(1-)=\mathcal{P}''(1-)\mu^{n-1}(\mu^n-1)/(\mu-1),\quad \mu\neq1,$$

$$\mathcal{P}''_n(1-)=n\mathcal{P}''(1-),\quad \mu=1,$$

$$\mathbf{D}X_n=\mathcal{P}''_n(1-)+\mu^n-\mu^{2n}=n\mathbf{D}X_1,\ \mu=1,$$

$$\mathbf{D}X_n=\mu^n(\mu^n-1)\cdot\mathbf{D}X_1/(\mu^2-\mu),\quad \mu\neq1.$$

Отметим, что при $\mu<1$ из неравенства (см. § 8)

$$\mathrm{P}(\omega : X_n(\omega) \geqslant t) \leqslant \mathrm{M}X_n t^{-1} = \mu^n t^{-1}$$

видно, что сходимость $X_n \xrightarrow{p} 0$ весьма «быстрая». При $\mu=1$ $\mathrm{M}X_n = 1$, дисперсия растет линейно по n, при $\mu>1$ $\mathrm{M}X_n=\mu^n$, $\mathrm{D}X_n$ имеет порядок квадрата математического ожидания. В обоих случаях использование неравенства Чебышева

$$\mathrm{P}(\omega : |X_n(\omega) - \mathrm{M}X_n| > t(\mathrm{D}X_n)^{1/2}) \leqslant t^{-2}$$

не позволяет оценить характер перемещения с ростом n вероятностной массы с законом распределения сл. в. X_n.]

14.12. Задача. Пусть вероятностная п. ф. $\mathscr{P}(s)$ дробно-линейна: $\mathscr{P}(s)=(a+bs)/(c+ds)$. Показать, что распределение вероятностей с п. ф. $\mathscr{P}(s)$ можно представить в виде

$$p_n = \alpha\beta^{n-1}, \quad n \geqslant 1, \quad p_0 = (1-\alpha-\beta)/(1-\beta), \quad \alpha, \ \beta \geqslant 0, \ \alpha+\beta<1,$$

так что $\mathscr{P}(s) = 1-\alpha/(1-\beta)+\alpha s/(1-\beta s)$. Вывести представление

$$1-\mathscr{P}(s) = (((1-s)\mathscr{P}'(1))^{-1} + \mathscr{P}''(1)/(2\mathscr{P}'(1))^2)^{-1}.$$

Выразить α и β через $\mathscr{P}'(1)$, $\mathscr{P}''(1)$.

[Пусть $\mathscr{P}(s)=(a+bs)/(c+ds)$. Так как из четырех коэффициентов a, b, c, d один, очевидно, можно взять любым фиксированным ($\neq 0$), то положим $c=1$. При $|ds|<1$, разлагая $\mathscr{P}(s)$ в ряд, получим для его коэффициентов выражения

$$p_n = (b-ad)(-d)^{n-1}, \quad n \geqslant 1,$$

откуда видно, что $0 \leqslant -d < 1$, либо $b-ad=0$ и $\mathscr{P}(s)=a$. Отбрасывая случай распределения, сосредоточенного в нуле, положим $\beta = -d$, $0 \leqslant \beta < 1$, $\alpha = b+a\beta > 0$, так что $p_n = \alpha\beta^{n-1}$ при $n \geqslant 1$. Записав условие $\mathscr{P}(1)=1$, получим систему для определения a и b:

$$a+b=1-\beta, \quad a\beta+b=\alpha,$$

откуда

$$p_0 = a = 1-\alpha/(1-\beta), \quad \mathscr{P}(s) = 1-\alpha/(1-\beta)+\alpha s/(1-\beta s).$$

Элементарные преобразования приводят к формуле

$$1-\mathscr{P}(s) = ((1-\beta)^2\alpha^{-1}(1-s)^{-1} + \beta(1-\beta)\alpha^{-1})^{-1}.$$

Кроме того,

$$\mathscr{P}'(s) = \alpha(1-\beta s)^{-2}, \quad \mathscr{P}''(s) = 2\alpha\beta(1-\beta s)^{-3},$$

$$\mathscr{P}'(1) = \alpha(1-\beta)^{-2}, \quad \mathscr{P}''(1) = 2\alpha\beta(1-\beta)^{-3},$$

что и приводит к требуемому представлению для $1-\mathscr{P}(s)$.

Наконец, разделив $\mathcal{P}''(1)$ на $\mathcal{P}'(1)$, найдем β и α:

$$\beta = \mathcal{P}''(1)/(2\mathcal{P}'(1) + \mathcal{P}''(1)),$$

$$\alpha = \mathcal{P}'(1)(1-\beta)^2 = 4\mathcal{P}'(1)^3/(2\mathcal{P}'(1) + \mathcal{P}''(1))^2.]$$

14.13. Задача. Для ветвящегося процесса с дробно-линейной п. ф., начинающегося с одной частицы, найти распределение числа частиц в n-м поколении.

Так как суперпозиция дробно-линейных функций является, очевидно, дробно-линейной, то для п. ф. $\mathcal{P}_n(s)$ числа потомков в n-м поколении получаем из 14.12

$$\mathcal{P}_n(s) = 1 - \alpha_n/(1-\beta_n) + \alpha_n s/(1-\beta_n s),$$

$$\alpha_n = 4\mathcal{P}_n'(1)^3/(2\mathcal{P}_n'(1) + \mathcal{P}_n''(1))^2, \quad \beta_n = \mathcal{P}_n''(1)/(2\mathcal{P}_n'(1) + \mathcal{P}_n''(1)).$$

Подставляя сюда выражение для $\mathcal{P}_n'(1)$, $\mathcal{P}_n''(1)$ из 14.11, приходим к следующим формулам для α_n, β_n:

$$\alpha_n = (2/(2 + n\mathcal{P}''(1)))^2, \quad \beta_n = n\mathcal{P}''(1)/(2 + n\mathcal{P}''(1))$$

при $\mu = 1$, а при $\mu \neq 1$ имеем

$$\alpha_n = \mu^n((p-1)/(p-\mu^n))^2, \quad \beta_n = (1-\mu^n)/(p-\mu^n),$$

где $p = 1 - \mu(\mu-1)2\mathcal{P}''(1)^{-1}$ (p представляет собой, как нетрудно проверить, меньший корень уравнения $\mathcal{P}(s) = s$). Распределение вероятностей с п. ф. $\mathcal{P}_n(s)$ получаем из 14.12.]

14.14. Задача. Рассмотрим ветвящийся процесс с дробно-линейной п. ф., начинающийся из точки 1. Предположим, что $\mu \leqslant 1$ и обозначим $T(\omega)$ время до вырождения процесса (которое конечно с вероятностью 1). Найти асимптотику хвостов его распределения $Q_t = \mathbf{P}(\omega : T(\omega) > t)$, $t \to \infty$.

[При $t > 0$ целом имеем

$$Q_t = \mathbf{P}(\omega : X_t(\omega) > 0) = 1 - \mathbf{P}(X_t(\omega) = 0) = 1 - \mathcal{P}_t(0).$$

Подставляя в $\mathcal{P}_t(0) = 1 - \alpha_t/(1-\beta_t)$ значения α_t, β_t из 14.13, находим при $\mu = 1$ и $\mu < 1$ соответственно

$$Q_t = \frac{4(2 + t\mathcal{P}''(1))}{(2 + t\mathcal{P}''(1))^2 \cdot 2} \sim \frac{2}{t\mathcal{P}''(1)}, \quad t \to \infty,$$

$$Q_t = \frac{\mu^t(p-1)^2(p-\mu^t)}{(p-\mu^t)^2(p-1)} \sim \frac{1-p}{p}\mu^t, \quad t \to \infty.$$

Отсюда, в частности, вытекает, что $\mathbf{M}T < \infty$ при $\mu < 1$ и $\mathbf{M}T = \infty$ при $\mu = 1$.]

14.15. Задача. Рассмотрим процесс Гальтона—Ватсона при условиях $\mu = \mathcal{P}'(1-) < 1$, $\mathcal{P}''(1-) < \infty$. Вывести для вероятностей $Q_t = \mathbf{P}(\omega : X_t(\omega) > 0)$ соотношение, используя формулу Тейлора:

$$Q_{t+1} = \mathcal{P}'(1-)Q_t - \frac{1}{2}\mathcal{P}''(\theta_t)Q_t^2, \quad 0 \leqslant \theta_t \leqslant 1.$$

Получить отсюда асимптотику

$$Q_t \sim \theta\mu^t, \quad \theta = \text{const} > 0.$$

Используя разложение по формуле Тейлора, получаем

$$Q_{t+1} = 1 - \mathcal{P}_{t+1}(0) = 1 - \mathcal{P}(\mathcal{P}_t(0)) = \mathcal{P}'(1-)Q_t - \frac{1}{2}\mathcal{P}''(\theta_t)Q_t^2.$$

Поделив это соотношение на μ^{t+1} и положив $R_t = Q_t\mu^{-t}$, $a_t = \mathcal{P}''(\theta_t)/2\mu$, запишем

$$R_{t+1} = R_t(1 - a_tQ_t), \quad R_n/R_1 = \prod_{t=1}^{n-1}(1 - a_tQ_t).$$

Так как $Q_{t+1} \leqslant \mu Q_t \leqslant \ldots \leqslant \mu^t Q_1$, а последовательность a_t, $t = 1, 2, \ldots$, ограничена, то

$$\lim_{n\to\infty} \ln \prod_{i=1}^{n}(1 - a_tQ_t) = \lim_{n\to\infty} \sum_{t=1}^{n} \ln(1 - a_tQ_t) =$$

$$= \lim_{n\to\infty} \sum_{t=1}^{n} (-a_tQ_t + O(1)a_t^2Q_t^2) < \infty,$$

т. е. существует конечный положительный предел последовательности R_n, $n = 1, 2, \ldots$, что и требовалось показать.

Для целей дальнейшего использования заметим, что из проведенных рассуждений вытекает следующее уточнение результата:

$$Q_n = \theta\mu^n + O(1)\mu^{2n}, \quad n \to \infty.$$

В самом деле, полагая $\theta = \lim\limits_{n\to\infty} R_n$, имеем

$$\ln R_n - \ln\theta = \sum_{t=n}^{\infty} \ln(1 - a_tQ_t) = O(\mu^n),$$

$$R_n = \theta\exp(O(\mu^n)) = \theta(1 + O(\mu^n)), \quad n \to \infty.]$$

14.16. З а д а ч а. Рассмотрим рекуррентное уравнение

$$Q_{t+1} = Q_t - aQ_t^2, \quad t = 1, 2, \ldots, \quad a > 0.$$

Предположим, что $Q_t \to 0$ при $t \to \infty$. Показать, что

$$Q_t^{-1} \sim at, \quad t \to \infty.$$

[$Q_{t+1}^{-1} = (Q_t - aQ_t^2)^{-1} = Q_t^{-1}(1 - aQ_t)^{-1}$, так что при $t \to \infty$

$$Q_{t+1}^{-1} = Q_t^{-1} + a + O(1)Q_t.$$

Суммируя, получаем

$$Q_{t+1}^{-1}-Q_1^{-1}=at+O(1)\sum_{i=1}^{t} Q_i.$$

Ввиду того что $Q_t \to 0$ при $t\to\infty$, остаточный член в правой части есть $o(t)$ и, следовательно, $Q_{t+1}^{-1}\sim at$.]

14.17. Задача. Рассмотрим процесс Гальтона—Ватсона при условиях $\mu=\mathcal{P}'(1-)=1$, $\mathcal{P}'''(1-)<\infty$. Вывести для вероятностей $Q_t=\mathbf{P}(\omega:X_t(\omega)>0)$ соотношение, используя формулу Тейлора:

$$Q_{t+1}=Q_t-\frac{1}{2}\mathcal{P}''(1-)Q_t^2+\frac{1}{3!}\mathcal{P}'''(\theta_t)Q_t^3,\quad 0<\theta_t<1.$$

Получить отсюда асимптотику

$$Q_t\sim 2(\mathcal{P}''(1-)t)^{-1},\quad t\to\infty.$$

[Рекуррентные соотношения для Q_t выводятся так же, как и в 14.15, а асимптотика устанавливается аналогично 14.16.]

Для процесса Гальтона—Ватсона $X_k(\omega)$, $k=0, 1, 2$, определим *условный процесс* при условии события $\Lambda_n=\{\omega:X_n(\omega)>0\}$, означающего, что процесс не выродился до момента времени n:

$$Y_k^{(n)}=(X_k\,|\,\Lambda_n),\quad k=0,\ 1,\ \ldots,\ n,$$

$$p_{Y_0,Y_1,\ldots,Y_n}(i_0,\ i_1,\ \ldots,\ i_n)=\mathbf{P}(\{\omega:X_k(\omega)=i_k,\ k=0,\ 1,\ \ldots,\ n\}\,|\,\Lambda_n).$$

14.18. Задача. Получить предельную теорему для $Y_n^{(n)}=(X_n\,|\,\Lambda_n)$, $n=1, 2,\ldots$, где X_n — ветвящийся процесс с дробно-линейной п. ф. и $\mu<1$, $\Lambda_n=\{\omega:X_n(\omega)>0\}$.
[Из 14.13, 14.14 имеем при $n\to\infty$

$$p_{Y_n}(j)=\mathbf{P}(\omega:X_n(\omega)=j)/\mathbf{P}(\Lambda_n)\to(1-1/p)(1/p)^{j-1},\quad j=1,\ 2,\ \ldots\ .]$$

14.19. Задача. Получить предельную теорему для $Y_n^{(n)}/n$, где $Y_n^{(n)}=(X_n|\Lambda_n)$, X_n — ветвящийся процесс с дробно-линейной п. ф. и $\mu=1$, $\Lambda_n=\{\omega:X_n(\omega)>0\}$.
[Из 14.13 получаем

$$p_{Y_n^{(n)}}(j)=p_{X_n}(j)/\mathbf{P}(\Lambda_n)=\frac{2}{2+n\mathcal{P}''(1)}\left(1-\frac{2}{2+n\mathcal{P}''(1)}\right)^{j-1},\quad j>0.$$

При $n\to\infty$ равномерно по $j=o(n^2)$ выполняются соотношения (локальная предельная теорема):

$$p_{Y_n^{(n)}}(j)\sim 2(n\mathcal{P}''(1))^{-1}\exp(-2j/(n\mathcal{P}''(1))).$$

Переходя к суммам, получаем интегральную предельную теорему

$$F_{Y_n^{(n)}/n}(x)\to 1-\exp(-2x/\mathcal{P}''(1)),\quad x\geqslant 0,\ n\to\infty.]$$

14.20. З а д а ч а. Решить задачи 14.18, 14.19, используя метод производящих функций (13.13) и преобразований Лапласа (13.28) соответственно.

[П. ф. сл. в. $Y_n^{(n)}$ равна

$$\sum_{k=1}^{\infty} s^k \mathbf{P}(\omega : X_n(\omega)=k)/\mathbf{P}(\omega : X_n(\omega)>0)=(\mathcal{P}_n(s)-\mathcal{P}_n(0))/(1-\mathcal{P}_n(0)).$$

Подставляя сюда выражение для $\mathcal{P}_n(s)$ из 14.13, получим

$$\mathbf{M}s^{Y_n^{(n)}}=(1-\beta_n)s/(1-\beta_n s).$$

При $\mu<1$ $\beta_n\to 1/p$, $n\to\infty$ и п. ф. сл. в. $Y_n^{(n)}$ сходятся к п. ф., сдвинутого на единицу геометрического распределения из 14.18. При $\mu=1$ рассмотрим преобразование Лапласа сл. в. $Y_n^{(n)}/n$. Учитывая, что $1-\beta_n\sim 2(n\mathcal{P}''(1))^{-1}$, получаем

$$\mathbf{M}e^{-\lambda Y_n^{(n)}/n}=e^{-\lambda/n}\left(\frac{1}{1-\beta_n}-\frac{\beta_n}{1-\beta_n}e^{-\lambda/n}\right)^{-1}\to(1+\lambda\mathcal{P}''(1)/2)^{-1},$$

что представляет собой преобразование Лапласа экспоненциального распределения (см. 13.15).]

14.21. З а д а ч а. В условиях 14.15 показать, что $Y_n^{(n)}=(X_n\mid\Lambda_n)$ имеет предельное распределение, рассмотрев предел соответствующих п. ф. Показать, **что предельная п. ф.**

$$\mathcal{P}^*(s)=\lim_{n\to\infty}\mathbf{M}s^{Y_n^{(n)}},\quad 0\leqslant s\leqslant 1,$$

удовлетворяет уравнению

$$1-\mathcal{P}^*(\mathcal{P}(s))=\mu(1-\mathcal{P}^*(s))$$

и потому отвечает собственному распределению ($\mathcal{P}^*(1)=1$).

[Рассуждения подобны проведенным в 14.15. П. ф. сл. в. $Y_n^{(n)}$ равна (см. 14.20)

$$(\mathcal{P}_n(s)-\mathcal{P}_n(0))/(1-\mathcal{P}_n(0))=1-(1-\mathcal{P}_n(s))/(1-\mathcal{P}_n(0)).$$

Полагая $R_n(s)=(1-\mathcal{P}_n(s))/(1-\mathcal{P}_n(0))$ и раскладывая $\mathcal{P}(s)$ по формуле Тейлора в точке $s=1$, запишем

$$1-\mathcal{P}(\mathcal{P}_n(s))=\mathcal{P}'(1)(1-\mathcal{P}_n(s))-\frac{1}{2}\mathcal{P}''(\theta_{n,s})(1-\mathcal{P}_n(s))^2,$$

$$R_{n+1}(s)=\mu\frac{1-\mathcal{P}_{n+1}(0)}{1-\mathcal{P}_n(0)}R_n(s)(1-O(1-\mathcal{P}_n(s))).$$

Из 14.15 имеем при $n\to\infty$

$$1-\mathcal{P}_n(s)\leqslant 1-\mathcal{P}_n(0)=\theta\mu^n+O(\mu^{2n}),$$

$$(1-\mathcal{P}_{n+1}(0))/(1-\mathcal{P}_n(0))=\mu+O(\mu^n),$$

так что полученное выше соотношение переписывается в виде

$$R_{n+1}(s)=R_n(s)(1+O(\mu^n)),$$

откуда и следует сходимость $R_n(s)$ к ненулевому пределу (ср. 14.15).

Точно так же получаем соотношение

$$(1-\mathcal{P}_n(\mathcal{P}(s)))/(1-\mathcal{P}_n(0))=\mu(1-\mathcal{P}_n(s))/(1-\mathcal{P}_n(0))(1+O(\mu^n)),$$

или

$$R_n(\mathcal{P}(s))=\mu R_n(s)(1+O(\mu^n)).$$

Переходя к пределу, приходим к уравнению

$$1-\mathcal{P}^*(\mathcal{P}(s))=\mu(1-\mathcal{P}^*(s)).$$

Полагая $s=1$, убеждаемся, что $\mathcal{P}^*(1)=1$.]

14.22. Задача. В условиях 14.17 показать, что $Y_n^{(n)}/n$, где $Y_n^{(n)}=(X_n\,|\,\Lambda_n)$ имеет экспоненциальное предельное распределение.

[Рассуждения подобны проведенным в 14.17. По формуле Тейлора имеем, полагая $R_n(s)=1-\mathcal{P}_n(s)$:

$$R_{n+1}(s)=R_n(s)\left(1-\frac{1}{2}\mathcal{P}''(1-)R_n(s)+O(R_n(s)^2)\right),$$

$$R_{n+1}(s)^{-1}=R_n(s)^{-1}\left(1+\frac{1}{2}\mathcal{P}''(1-)R_n(s)+O(R_n(s)^2\right),$$

$$R_{n+1}(s)^{-1}=R_n(s)^{-1}+\frac{1}{2}\mathcal{P}''(1-)+O(R_n(0)),\quad n\to\infty,$$

где мы учли, что $R_n(s)\leqslant R_n(0)$. Как и в 14.16, отсюда получаем

$$R_n(s)^{-1}=(1-s)^{-1}+\frac{1}{2}\mathcal{P}''(1-)n+o(n),\ n\to\infty.$$

Подставляя $s=e^{-\lambda/n}$, находим

$$R_n(e^{-\lambda/n})^{-1}=n\lambda^{-1}+\frac{1}{2}\mathcal{P}''(1-)n+o(n),\ n\to\infty.$$

Наконец (см. 14.20 и 14.19),

$$\mathbf{M}\exp(-\lambda Y_n^{(n)}/n)=1-R_n(e^{-\lambda/n})/(1-\mathcal{P}_n(0))\to$$

$$\to 1-((2/\mathcal{P}''(1))\lambda^{-1}+1)^{-1}=(1+(\mathcal{P}''(1)/2)\lambda)^{-1},\ n\to\infty.]$$

14.23. Задача. Рассмотрим процесс Гальтона—Ватсона $X_n^{(n)}$, $n=0, 1, 2, \ldots$, при $\mu=\mathcal{P}'(1-)=1$, $\mathcal{P}'''(1-)<\infty$, $\gamma=\mathcal{P}''(1-)/2$ и предположим, что он начался с «большого» числа частиц: $X_0^{(n)}=[n\gamma x]$, $x>0$. Показать, что последовательность сл. в. $X_n^{(n)}/(n\gamma)$, $n=1, 2, \ldots$, сходится по распределению.

Процесс Гальтона—Ватсона, начинающийся с k частиц, складывается из k независимых процессов, начинающихся с одной частицы, так что

$$X_n^{(n)} \overset{\mathscr{D}}{=} \sum_{i=1}^{[n\gamma x]} X_{ni},$$

где X_{ni}, $n=0, 1, 2, \ldots$, — процесс, начинающийся с одной частицы, а совокупность этих процессов при $i=1, 2, \ldots$ независима. Поскольку $\mathbf{P}(\omega : X_{ni}(\omega)>0) \sim 1/(n\gamma)$, $n\to\infty$, то по теореме Пуассона число N_n ненулевых слагаемых в сумме имеет в пределе распределение

$$p_{N_n}(k) \to (1/k!)\, x^k e^{-x}, \quad k=0, 1, 2, \ldots .$$

Поскольку сл. в. $X_{ni}/(n\gamma)$ при условии $X_{ni}>0$ имеют в пределе стандартное экспоненциальное распределение, то отсюда легко вывести, что предельное распределение сл. в. $X_n^{(n)}/(n\gamma)$ совпадает с распределением суммы пуассоновского числа независимых экспоненциальных слагаемых. Легко выписать преобразование Лапласа этого распределения, однако формальные выкладки, основанные на 14.22, приводят к требуемому результату проще:

$$\mathbf{M}\exp(-\lambda X_n^{(n)}/(n\gamma)) = (\mathscr{P}_n(e^{-\lambda/(n\gamma)}))^{[n\gamma x]} = (1-R_n(e^{-\lambda/n\gamma}))^{[n\gamma x]} =$$
$$= (1-(1+\lambda^{-1}+o(1)^{-1}(\gamma n)^{-1})^{[n\gamma x]} \to e^{-\lambda x/(1+\lambda)}).$$

14.24. Задача. Сл. в. N, X_1, X_2, $\ldots$ независимы в совокупности, сл. в N распределена по Пуассону с параметром β, сл. в. X_i, $i=1, 2, \ldots$, имеют стандартные экспоненциальные распределения. Найти преобразование Лапласа и распределение вероятностей суммы $S_N=X_1+\ldots+X_N$ ($S_N=0$ при $N=0$).

[По формуле 9.3

$$\mathbf{M}e^{-\lambda S_N} = \sum_{n=0}^{\infty} \mathbf{M}(e^{-\lambda S_N} I_{(\omega : N(\omega)=n)}).$$

При $n>0$ запишем слагаемые суммы в виде

$$\mathbf{M}(e^{-\lambda S_n} I_{(\omega : N(\omega)=n)}) = \mathbf{M}e^{-\lambda S_n}\cdot \mathbf{M}I_{(\omega : N(\omega)=n)} =$$
$$= (\mathbf{M}e^{-\lambda X_1})^n \cdot \mathbf{P}(\omega : N(\omega)=n) = (1+\lambda)^{-n}(\beta^n/n!)\, e^{-\beta},$$

откуда получаем

$$\mathbf{M}e^{-\lambda S_N} = \sum_{n=0}^{\infty} (1+\lambda)^{-n}(\beta^n/n!)\, e^{-\beta} = e^{-\beta\lambda/(1+\lambda)}$$

Аналогично находим распределение вероятностей:

$$\mathbf{P}(\omega : S_n(\omega) \in I_{a,b}) = \sum_{n=0}^{\infty} \mathbf{P}(\omega : S_n(\omega) \in I_{a,b}) \cdot (\beta^n/n!)\, e^{-\beta}.$$

Сл. в. S_N имеет гамма-плотность (см. 4.19)

$$f_{S_n}(y) = (1/(n-1)!)\, y^{n-1} e^{-y}, \quad y > 0.$$

Таким образом, ф. р. $F(y)$ сл. в. S_N имеет скачок величины $\mathbf{P}(\omega : N(\omega) = 0) = e^{-\beta}$ в нуле, при $y > 0$ она дифференцируема, а ее производная равна

$$\sum_{n=1}^{\infty} \frac{y^{n-1}\beta^n}{(n-1)!\, n!} e^{-y-\beta} \equiv f(y;\ \beta). \Big]$$

14.25.* Пример (условный ветвящийся процесс, $\mu = 1$). Совместное распределение значений условного ветвящегося процесса $Y_l^{(n)} = (X_l \mid \Lambda_n)$, $\Lambda_n = \{\omega : X_n(\omega) > 0\}$, $Y_0^{(n)} = 1$, в точках $0 < n_1 < \ldots < n_k < n$ представимо в виде

$$p_{Y_{n(1)}^{(n)}, \ldots, Y_{n(k)}^{(n)}}(i_{n_1}, \ldots, i_{n_k}) = \sum p_{1 i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n} Q_n^{-1},$$

где индексы $i_{n_1}, \ldots, i_{n_k} > 0$ в сумме справа фиксированы, а по всем остальным индексам идет суммирование по положительным (>0) значениям, Q_n — вероятность невырождения в течение n шагов процесса, начавшегося с одной частицы. Обозначим $Q_n^{(j)}$, $j = 1, 2, \ldots$, вероятность невырождения за время n процесса, начавшегося с j частиц, $Q_n^{(1)} \equiv Q_n$.

Напомним, что вероятность перехода за m шагов определяется как

$$p_{ij}^{(m)} = \sum_{(i_1, \ldots, i_{m-1})} p_{i i_1} p_{i_1 i_2} \ldots p_{i_{m-1} j},$$

где суммирование ведется по всем $i_1, \ldots, i_{m-1} \geqslant 0$. Так как $p_{0k} = 0$ при $k > 0$, то для $i, j > 0$ в сумме, определяющей $p_{ij}^{(m)}$, можно брать только положительные (>0) значения индексов суммирования. С учетом этого замечания совместное распределение сл. в. $Y_{n_1}, \ldots, Y_{n_k}$ примет вид ($n_j \equiv n(j)$):

$$p_{1 i_{n(1)}}^{(n(1))} p_{i_{n(1)} i_{n(2)}}^{(n(2)-n(1))} \ldots p_{i_{n(k-1)} i_{n(k)}}^{(n(k)-n(k-1))} Q_{n-n(k)}^{(i_{n(k)})} / Q_n.$$

Положим при $0 < n_1 < n_2 \leqslant n$, $i, j > 0$

$$p(i, n_1, j, n_2) = p_{ij}^{(n(2)-n(1))} Q_{n-n_2}^{(j)} / Q_{n-n_1}^{(i)}.$$

Так как, очевидно,

$$\sum_{(j>0)} p_{ij}^{(n(2)-n(1))} Q_{n-n_2}^{(j)} = Q_{n-n_1}^{(i)},$$

то $p(i, n_1, j, n_2)$ представляет собой распределение вероятностей на множестве натуральных чисел. Легко видеть, что совместное распределение можно переписать в виде

$$p_{Y_{n(1)}^{(n)}, \ldots, Y_{n(k)}^{(n)}}(i_{n_1}, \ldots, i_{n_k}) = \prod_{j=0}^{k-1} p(i_{n_j}, n_j, i_{n_{j+1}}, n_{j+1}),$$

где полагаем $n_0=0$, $i_{n_0}=1$, т. е. условный процесс $Y_l^{(n)}$, $l=0, 1, \ldots$ $\ldots, n$ — неоднородная марковская цепь с переходной функцией $p(i, n_1, j, n_2)$.

Далее предположим, что $\mu=\mathcal{P}'(1-)=1$, $\mathcal{P}'''(1-)<\infty$, обозначим $\mathcal{P}''(1-)/2=\gamma$ и введем процесс с непрерывным временем

$$Z_t^{(n)} = Y_{[nt]}^{(n)}/(\gamma n), \quad 0 \leqslant t \leqslant 1,$$

где $[a]$ есть целая часть числа a. Покажем, что конечномерные распределения процесса $Z_t^{(n)}$ слабо сходятся при $n\to\infty$ к распределениям некоторого марковского процесса Z_t.

Начнем с одномерных распределений. При $0<t\leqslant 1$, $j>0$ запишем

$$\mathbf{P}(\omega : Z_t^{(n)}(\omega) = j/(\gamma n)) = \mathbf{P}(\omega : Y_{[nt]}^{(n)}(\omega) = j) = p(1, 0, j, [nt]) =$$
$$= p_{1j}^{([nt])} Q_{n-[nt]}^{(j)}/Q_n = (p_{1j}^{([nt])}/Q_{[nt]})\, Q_{n-[nt]}^{(j)}\,(Q_{[nt]}/Q_n).$$

Заметим, что

$$p_{1j}^{(m)}/Q_m = \mathbf{P}(\omega : Y_m^{(m)}(\omega) = j),$$

а потому из 14.22 получаем при $n\to\infty$ и любых $0<a<b$:

$$\sum_{(j \in I\gamma an, \gamma bn)} p_{1j}^{([nt])}/Q_{[nt]} \to \int_a^b t^{-1} e^{-x/t}\, dx.$$

Из асимптотики вероятностей невырождения 14.17 имеем

$$Q_{[nt]}/Q_n \to t^{-1}, \quad n\to\infty.$$

Наконец, так как $1-Q_m^{(j)}$ есть вероятность вырождения до момента m включительно процесса, начавшегося с j частиц, а процессы размножения каждой из j начальных частиц независимы, то

$$1-Q_m^{(j)} = (1-Q_m)^j = (1-1/(\gamma m) + o(1/m))^j \sim e^{-j/(\gamma m)}$$

равномерно по $j=O(m)$, $m\to\infty$.

Покажем, что при $n \to \infty$

$$\mathbf{P}(\omega : Z_t^{(n)}(\omega) \in I_{a,b}) \to \int_a^b t^2 e^{-x/t}(1 - e^{-x/(1-t)})\,dx.$$

Рассмотрим произвольное разбиение $a = x_0 < x_1 < \ldots < x_k = b$ промежутка $(a, b]$ и запишем

$$\mathbf{P}(\omega : Z_t^{(n)}(\omega) \in (a, b]) = \sum_{i=1}^{k} \mathbf{P}(\omega : Z_t^{(n)}(\omega) \in (x_{i-1}, x_i]) =$$

$$= \sum_{i=1}^{k} \sum_{\{j \in (n\gamma x_{i-1}, n\gamma x_i]\}} (p_{1j}^{([nt])}/Q_{[nt]})\, Q_{n-[nt]}^{(j)}\,(Q_{[nt]}/Q_n).$$

Так как вероятности $Q_{n-[nt]}^{(j)}$ возрастают по j, то, заменяя их при $j \in (n\gamma x_{i-1}, n\gamma x_i]$ на значения в крайних точках промежутка, получим для них оценки снизу и сверху. Переходя к пределу по n, получим

$$\varliminf_{n\to\infty} \mathbf{P}(\omega : Z_t^{(n)}(\omega) \in (x_{i-1}, x_i]) \geqslant \int_{x_{i-1}}^{x_i} t^{-2} e^{-x/t}\,dx \cdot (1 - e^{-x_i/(1-t)}).$$

Суммируя по $i = 1, \ldots, k$ и переходя к пределу по последовательности измельчающихся разбиений, имеем

$$\varliminf_{n\to\infty} \mathbf{P}(\omega : Z_t^{(n)}(\omega) \in (a, b]) \geqslant \int_a^b t^{-2} e^{-x/t}(1 - e^{-x/(1-t)})\,dx.$$

Такая же оценка имеет место для верхнего предела, что и доказывает высказанное утверждение.

Рассмотрим условное распределение

$$\mathbf{P}(\{\omega : Z_{t_2}^{(n)}(\omega) \in I_{a,b}\} \mid \{\omega : Z_{t_1}^{(n)}(\omega) = x\}) =$$

$$= \mathbf{P}(\{\omega : Y_{[nt_2]}^{(n)}(\omega) \in I_{an\gamma, bn\gamma}\} \mid \{\omega : Y_{[nt_1]}^{(n)}(\omega) = x\gamma n\}),$$

где $0 < a < b$, $0 < t_1 < t_2 \leqslant 1$, $x > 0$, $x\gamma n$ — целое. Оно получается суммированием по $j \in I_{an\gamma,\, bn\gamma}$ переходных вероятностей

$$p(i, [nt_1], j, [nt_2]) = p_{ij}^{([nt_2]-[nt_1])}\,(Q_{n-[nt_2]}^{(j)}/Q_{n-[nt_1]}^{(i)}),$$

где $i = x\gamma n$. Второй множитель в правой части имеет асимптотику

$$(1 - \exp(-j/\gamma n(1-t_2))))/(1 - \exp(-i/\gamma n(1-t_1)))),$$

что же касается первого множителя, то он представляет собой вероятность того, что ветвящийся процесс, начинающийся с $i = x\gamma n$ частиц, попадает по прошествии времени $[nt_2] - [nt_1]$ в состояние j ($j \in I_{an\gamma,\, bn\gamma}$). Применяя 14.23, 14.24, получаем для любого действительного $x>0$, $i=[x\gamma n]$, что

$$\sum_{\{j \in I_{an\gamma, bn\gamma}\}} p_{ij}^{([nt_2]-[nt_1])} \to \int\limits_{a/(t_2-t_1)}^{b/(t_2-t_1)} f(y;\ x/(t_2-t_1))\, dy.$$

Вводя, как и выше, разбиение промежутка $I_{a,b}$, приходим к предельной теореме

$$\mathbf{P}(\{\omega : Z_{t_2}^{(n)}(\omega) \in I_{a,b}\} \mid \{\omega : Z_{t_1}^{(n)}(\omega) = [x\gamma n]/(\gamma n)\}) \to$$

$$\to \int\limits_{a/(t_2-t_1)}^{b/(t_2-t_1)} f(y;\ x/(t_2-t_1)) \frac{1-\exp(-y(t_2-t_1)/(1-t_2))}{1-\exp(-x/(1-t_1))}\, dy =$$

$$= \int\limits_a^b x(t_2-t_1)^{-2} e^{-\frac{x+y}{t_2-t_1}} \sum_{k=1}^{\infty} \frac{(xy/(t_2-t_1)^2)^{k-1}}{k!\,(k-1)!} \cdot \frac{1-e^{-y/(1-t_2)}}{1-e^{-x/(1-t_1)}}\, dy.$$

Обозначим подынтегральную функцию в последнем выражении $f(t_1, x, t_2, y)$, $0<t_1<t_2\leqslant 1$, $x, y>0$. Легко видеть, что если положить здесь формально $t_1=x=0$, то получим плотность предельного распределения сл. в. $Z_{t_2}^{(n)}$. Использованная в проведенных выше рассуждениях техника может быть применена к доказательству того, что пределом по распределению последовательности процессов $Z_t^{(n)}$, $0\leqslant t\leqslant 1$, является неоднородный марковский процесс Z_t, $0\leqslant t\leqslant 1$, начинающийся из нуля и имеющий *переходную плотность* $f(s, x, t, y)$, $0\leqslant s<t\leqslant 1$, $x, y\geqslant 0$, т. е. совместное распределение сл. в. $Z_{t_1}, Z_{t_2}, \ldots, Z_{t_k}$, $0\leqslant t_1<t_2<\ldots<t_k\leqslant 1$, задается плотностью

$$f(0, 0, t_1, x_1) f(t_1, x_1, t_2, x_2) \ldots f(t_{k-1}, x_{k-1}, t_k, x_k).$$

Ограничимся по этому поводу одним замечанием, полагая для простоты $k=2$. Из предыдущего вытекает, что

$$\mathbf{P}(\omega : Z_{t_1}^{(n)}(\omega) \in I_{a,b},\ Z_{t_2}^{(n)}(\omega) \in I_{c,d}) = \sum_{(i \in I_{an\gamma, bn\gamma})} p_{1i}^{([nt_1])} \times$$

$$\times Q_{n-[nt_1]}^{(i)} Q_n^{-1} \sum_{(j \in I_{cn\gamma, dn\gamma})} p_{ij}^{([nt_2]-[nt_1])} Q_{n-[nt_2]}^{(j)} / Q_{n-[nt_1]}^{(i)}.$$

Чтобы при переходе к пределу получить в правой части повторный интеграл

$$\int\limits_a^b f(0,\ 0,\ t_1,\ x_1) \Big(\int\limits_c^d f(t_1,\ x_1,\ t_2,\ x_2)\, dx_2 \Big) dx_1,$$

достаточно, например, установить, что внутренняя сумма сходится к внутреннему интегралу равномерно по i, таким, что $i/(\gamma n) \in I_{a,b}$,

$0<a<b$, фиксированы. Однако это утверждение легко доказать, формализуя намеченный в 14.23 прямой путь доказательства предельной теоремы (т.е. не использующий преобразование Лапласа) и замечая, что предельная теорема Пуассона выполняется равномерно по параметру распределения в ограниченном и отделенном от нуля интервале.

14.26.* Пример (условный ветвящийся процесс, $\mu<1$). Как видно из предыдущего примера, если при $\mu=1$ ветвящийся процесс до момента $n\gg 1$ не выродился, то его траектории на временном интервале $[\varepsilon n, n]$ принимают значения порядка n. Совсем иной характер имеют траектории условного процесса $Y_l^{(n)}=(X_l\mid\Lambda_n)$, $l\leqslant n$, $Y_0^{(n)}=1$, $\Lambda_n=\{\omega:X_n(\omega)>0\}$ при $\mu<1$. Как и выше, примем ограничение $\mathscr{P}''(1-)<\infty$. При любых $i_1, i_2, \ldots, i_k>0$, $k\leqslant n$, имеем

$$p_{Y_1^{(n)},\ldots,Y_k^{(n)}}(i_1, \ldots, i_k)=\mathbf{P}(\{\omega: X_j(\omega)=i_j,\quad j=1, \ldots, k\}\mid\Lambda_n)=$$

$$=p_{1i_1}p_{i_1i_2}\ldots p_{i_{k-1}i_k}\cdot(1-p_{i_k0}^{(n-k)})\cdot(1-p_{10}^{(n)})^{-1}.$$

Положим

$$q_m(j\mid i)=p_{ij}(1-p_{j0}^{(n-m)})/(1-p_{i0}^{(n-m+1)}),$$

так что

$$p_{Y_1^{(n)},\ldots,Y_k^{(n)}}(i_1, \ldots, i_n)=\prod_{m=1}^{k} q_m(i_m\mid i_{m-1}),\quad i_0=1.$$

При этом

$$\sum_{\{j>0\}} p_{ij}(1-p_{j0}^{(n-m)})=\sum_{\{j>0\}} p_{ij}-\sum_{\{j>0\}} p_{ij}p_{j0}^{(n-m)}=$$

$$=(1-p_{i0})-\Big(\sum_{\{j\}} p_{ij}p_{j0}^{(n-m)}-p_{i0}p_{00}^{(n-m)}\Big)=1-p_{i0}^{(n-m+1)},$$

т. е. $q_m(j\mid i)$, $i, j=1, 2, \ldots,$ — переходные вероятности, сл. в. $Y_0^{(n)}=1$, $Y_1^{(n)}, \ldots, Y_n^{(n)}$ образуют неоднородную марковскую цепь. Выпишем распределение сл. в. $Y_k^{(n)}$:

$$p_{Y_k^{(n)}}(i)=p_{1i}^{(k)}(1-p_{i0}^{(n-k)})/(1-p_{10}^{(n)}).$$

Рассмотрим *обращенную марковскую цепь* $Z_l^{(n)}=Y_{n-l}^{(n)}$, $l=0, 1, \ldots, n$. Как и выше, получаем

$$p_{Z_0^{(n)},\ldots,Z_k^{(n)}}(i_0, i_1, \ldots, i_k)=p_{1i_k}^{(n-k)}p_{i_ki_{k-1}}\ldots p_{i_1i_0}Q_n^{-1},$$

где $Q_n=1-p_{10}^{(n)}$. Полагая при $i, j>0$, таких, что $p_{1j}^{(n-m)}>0$, $p_{1i}^{(n-m+1)}>0$,

$$\widehat{q}_m(j\mid i)=p_{1j}^{(n-m)}p_{ji}/p_{1i}^{(n-m+1)},\quad \sum_{\{j>0\}}\widehat{q}_m(j\mid i)=1,$$

запишем совместное распределение в виде

$$p_{Z_0^{(n)},\ldots,Z_k^{(n)}}(i_0,\ i_1,\ \ldots,\ i_k)=(p_{1i_0}^{(n)}/Q_n)\cdot\prod_{m=1}^{k}\widehat{q}_m(i_m\,|\,i_{m-1}),$$

так что сл. в. $Z_l^{(n)}$, $l=0,\ 1,\ \ldots,\ n$, образуют неоднородную марковскую цепь с переходными вероятностями $\hat{q}(j\,|\,i)$ и начальным распределением $p_{1i}^{(n)}/Q_n$ (зависящим от n). Отметим равенство, связывающее переходные вероятности прямой и обращенной марковских цепей

$$p_{Y_{k+1}^{(n)}|Y_k^{(n)}}(j\,|\,i)\cdot p_{Y_k^{(n)}}(i)=p_{Y_k^{(n)}|Y_{k+1}^{(n)}}(i\,|\,j)\cdot p_{Y_{k+1}^{(n)}}(j).$$

Исследуем предельное поведение распределений прямой цепи при $n\to\infty$, воспользовавшись полученными ранее результатами (см. 14.15, 14.21):

$$Q_n=1-p_{10}^{\prime(n)}\sim\theta\mu^n,\quad n\to\infty,$$

$$p_{1i}^{(n)}/Q_n\to p_i^*,\quad n\to\infty,\ i=1,\ 2,\ \ldots,$$

где p_i^*, $i=1,\ 2,\ \ldots$, представляет собой распределение вероятностей с п. ф. $\mathscr{P}^*(s)$, удовлетворяющей уравнению

$$1-\mathscr{P}^*(\mathscr{P}(s))=\mu(1-\mathscr{P}^*(s)).$$

Отметим еще, что

$$1-p_{i0}^{(k)}=1-(p_{10}^{(k)})^i=(1-p_{10}^{(k)})(1+p_{10}^{(k)}+\ldots+(p_{10}^{(k)})^{i-1})$$

и если $k\to\infty$, то

$$1-p_{i0}^{(k)}\sim\theta\mu^k i,\quad i=1,\ 2,\ \ldots.$$

Легко видеть, что равномерно по m, таким, что $n-m\to\infty$, выполняются соотношения

$$q_m(j\,|\,i)\to jp_{ij}/(i\mu)\equiv q_{ij},\quad i,\ j=1,\ 2,\ \ldots,$$

причем поскольку

$$\sum_{\{j>0\}}jp_{ij}=(d/ds)\mathscr{P}^i(s)|_{s=1-}=i\mathscr{P}^{i-1}(1)\mathscr{P}'(1)=i\mu.$$

(т. е. математическое ожидание числа потомков в первом поколении от i частиц равно сумме математических ожиданий числа потомков от каждой из этих i частиц), то q_{ij}, i, $j=1,\ 2,\ \ldots$, представляют собой переходные вероятности. Введем марковскую цепь $Y_0=1$, Y_1, Y_2, ... с переходными вероятностями q_{ij}. Из предыдущего вытекает, что при любых k, $i_1,\ldots,i_k>0$

$$p_{Y_1^{(n)},\ldots,Y_k^{(n)}}(i_1,\ \ldots,\ i_k)\to p_{Y_1,\ldots,Y_k}(i_1,\ \ldots,\ i_k),\quad n\to\infty.$$

В таком случае говорят, что процесс $Y_l^{(n)}$, $l=1, 2, \ldots$, *сходится по распределению* к процессу Y_l, $l=1, 2, \ldots$. На самом деле в нашем случае имеет место более сильный результат. Из первоначального выражения для совместного распределения сл. в. $Y_1^{(n)}, \ldots, Y_k^{(n)}$ находим, что

$$p_{Y_1^{(n)},\ldots,Y_k^{(n)}}(i_1, \ldots, i_k) \sim p_{1i_1}p_{i_1i_2}\ldots p_{i_{k-1}i_k}\mu^{-k}i_k = p_{Y_1,\ldots,Y_k}(i_1, \ldots, i_k)$$

равномерно по k, $i_1, \ldots, i_{k-1}>0$, такие, что $n-k\to\infty$. Понятно, что изучение процесса $Y_l^{(n)}$, $l=0, 1, 2, \ldots$, на временном участке $[0, n-k_n]$, где $k_n\to\infty$ сколь угодно медленно, сводится к исследованию марковской цепи Y_l, $l=0, 1, 2, \ldots$.

В особенности интересно в связи с процессом $Y_l^{(n)}$ изучить асимптотическое поведение марковской цепи Y_l при $l\to\infty$. Следующий важный результат непосредственно вытекает из предыдущего:

$$p_{Y_k}(i) = p_{1i}^{(k)}\mu^{-k}i \to p_i^*\theta i, \quad k\to\infty, \ i\geqslant 1.$$

Проверим, что предельное распределение

$$q_i \equiv p_i^*\theta i, \quad i=1, 2, \ldots,$$

сл. в. Y_k, $k=1, 2, \ldots$, является собственным: $q_1+q_2+\ldots=1$. Так как

$$\sum_{i=1}^{\infty} ip_i^* = \lim_{s\to 1-}\frac{1-\mathcal{F}^*(s)}{1-s},$$

а функция под знаком предела монотонна, то достаточно установить, что для некоторой последовательности $s_n\to 1$, $s_n<1$, предел $(1-\mathcal{F}^*(s_n))/(1-s_n)$ равен $1/\theta$. Итерируя уравнение, которому удовлетворяет п. ф. $\mathcal{F}^*(s)$, получаем (напомним, что $\mathcal{F}_k(s) \equiv \mathcal{F} \times \times (\mathcal{F}(\ldots\mathcal{F}(s)\ldots))$ n раз)

$$1-\mathcal{F}^*(\mathcal{F}_n(s)) = \mu(1-\mathcal{F}^*(\mathcal{F}_{n-1}(s))) = \ldots = \mu^n(1-\mathcal{F}^*(s)).$$

Взяв здесь $s=0$ и учитывая, что $\mathcal{F}^*(0)=0$, $\mathcal{F}_n(0)\to 1$, находим

$$\lim_{n\to\infty}\frac{1-\mathcal{F}^*(\mathcal{F}_n(0))}{1-\mathcal{F}_n(0)} = \lim_{n\to\infty}\frac{\mu^n}{1-\mathcal{F}_n(0)} = \theta^{-1}.$$

Итак, для удаленного от начала момента времени распределение сл. в. Y_l приближается к распределению q_i, $i=1, 2, \ldots$.

Примечательным свойством распределения q_i, $i=1, 2, \ldots$, является то, что оно удовлетворяет системе уравнений

$$q_j = \sum_{(i)} q_i q_{ij}, \quad j=1, 2, \ldots.$$

Хотя это нетрудно проверить непосредственно (и такой вывод проделан ниже), гораздо более важно, что указанная система — следствие предельных соотношений

$$\lim_{k\to\infty} p_{Y_k}(i)=q_i, \quad i=1,\ 2,\ \ldots .$$

В самом деле, записав уравнения Колмогорова—Чепмена

$$q_{ij}^{(n+1)}=\sum_{\{k\}} q_{ik}^{(n)}q_{kj}, \quad i,\ j=1,\ 2,\ \ldots,$$

имеем

$$p_{Y_{n+1}}(j)=\sum_{\{i\}} q_{1i}q_{ij}^{(n+1)}=\sum_{\{i\}} q_{1i}\sum_{\{k\}} q_{ik}^{(n)}q_{kj}=\sum_{\{k\}} q_{kj}\sum_{\{i\}} q_{1i}q_{ik}^{(n)}=\sum_{\{k\}} p_{Y_n}(k)\,q_{kj}.$$

Переход к пределу приводит к требуемой системе уравнений. Для доказательства оборвем сумму по k на конечном числе членов и, переходя к пределу по n, получим неравенства

$$q_j\geqslant\sum_{\{k\}} q_kq_{kj},$$

где суммирование ведется уже по всем k. Предположив, что здесь имеет место строгое неравенство, просуммируем по j и придем к противоречию:

$$\sum_{\{j\}} q_j>\sum_{\{k\}} q_k\sum_{\{j\}} q_{kj}=\sum_{\{k\}} q_k.$$

Введем марковскую цепь Y_k^*, $k=0,\ 1,\ 2,\ \ldots$, с теми же переходными вероятностями q_{ij} и начальным распределением

$$q_i=p_{Y_0^*}(i), \quad i=1,\ 2,\ \ldots .$$

Как и выше, имеем

$$p_{Y_{n+1}^*}(j)=\sum_{\{k\}} p_{Y_n^*}(k)\,q_{kj},$$

откуда индукцией по n получаем, что

$$p_{Y_n^*}(j)=q_j, \quad j=1,\ 2,\ \ldots,$$

при всех $n=0,\ 1,\ 2,\ \ldots$. Отсюда следует, что

$$p_{Y_{t+1}^*,\ldots,Y_{t+m}^*}(i_1,\ \ldots,\ i_m)=q_{i_1}q_{i_1i_2}\ldots q_{i_{m-1}i_m},$$

т. е. распределение отрезка марковской цепи $Y_{t+1}^*,\ \ldots,\ Y_{t+m}^*$ не зависит от сдвига времени. Это свойство называется *стационарностью*. Любой случайный процесс, конечномерные распределения которого не зависят от сдвига времени (в пределах временного интервала, на котором процесс задан), называется *стационар-*

ным. В данном случае имеем дело со *стационарной марковской цепью*.

Для исходного процесса Y_k, $k=0, 1, 2, \ldots$, распределение вероятностей

$$p_{Y_{t+1},\ldots,Y_{t+m}}(i_1, \ldots, i_m)=p_{Y_{t+1}}(i_1)\, q_{i_1 i_2} \ldots q_{i_{m-1} i_m}$$

приближается при $t\to\infty$ к стационарному, так как $p_{Y_{t+1}}(i)\to q_i$ при $t\to\infty$. Говорят, что марковская цепь Y_t, $t=0, 1, 2, \ldots$, *входит в стационарный режим* при $t\to\infty$. Примечательно, что это свойство не зависит от начального состояния:

$$q_{ij}^{(k)}\to q_j\equiv p_j^*\theta j, \quad k\to\infty, \; j=1, 2, \ldots,$$

для любого $i=1, 2, \ldots$. В самом деле,

$$q_{ij}^{(k)}=\sum_{(i_1,\ldots,i_{k-1})} q_{ii_1}q_{i_1i_2}\ldots q_{i_{k-1}j}=$$

$$=\sum_{(i_1,\ldots,i_{k-1})} p_{ii_1}\ldots p_{i_{k-1}j}i^{-1}j\mu^{-k}=i^{-1}jp_{ij}^{(k)}\mu^{-k},$$

и требуемый результат вытекает из соотношения

$$p_{ij}^{(k)}/(1-p_{i0}^{(k)})\to p_j^*, \quad k\to\infty, \; i, j=1, 2, \ldots,$$

которое устанавливается в задаче 14.27, и $1-p_{i0}^{(k)}\sim\theta\mu^k i$ (см. выше).

Произвольное распределение вероятностей π_i, $i=1, 2, \ldots$, называется *стационарным распределением марковской цепи* с переходными вероятностями q_{ij}, $i, j=1, 2, \ldots$, если оно удовлетворяет системе линейных уравнений

$$\pi_j=\sum_{(i)}\pi_i q_{ij}, \quad j=1, 2, \ldots.$$

Отсюда, очевидно, имеем

$$\pi_j=\sum_{(i)}\pi_i q_{ij}^{(n)}, \quad j, n=1, 2, \ldots.$$

Если существуют пределы

$$\lim_{n\to\infty} q_{ij}^{(n)}=q_j, \quad i, j=1, 2, \ldots, \quad \sum_{(j)} q_j=1,$$

то, как и ранее, убеждаемся, что возможен переход к пределу при $n\to\infty$ под знаком суммы, откуда получаем равенства $\pi_j=q_j$, $j=1, 2, \ldots$, т. е. у марковской цепи имеется единственное стационарное распределение.

Мы уже отмечали, что к изучению марковской цепи Y_t, $t=0, 1, 2, \ldots$, сводится исследование предельного поведения распределений условного ветвящегося процесса $Y_l^{(n)}$ при $l \in [0, n-k_n]$, где $k_n \to \infty$ сколь угодно медленно. Обращенная марковская цепь $Z_l^{(n)} = Y_{n-l}^{(n)}$ позволяет исследовать эти же вопросы для $l \in [k_n, n]$. Рассуждения здесь подобны проведенным выше. Прежде всего имеем при $n \to \infty$, k, $i_1, \ldots, i_k > 0$:

$$p_{Z_0^{(n)}, \ldots, Z_k^{(n)}}(i_0, \ldots, i_k) \to p_{i_k}^* p_{i_k} p_{i_{k-1}} \ldots p_{i_1 i_0} \mu^{-k}.$$

Обозначим $\mathcal{I}$ подмножество таких $i>0$, что $p_i^* > 0$; для i, $j \in \mathcal{I}$ положим

$$\hat{q}_{ij} = p_j^* p_{ji} / (p_i^* \mu).$$

Проверим, что $\hat{q}_{ij}$, i, $j \in \mathcal{I}$, — переходные вероятности. Это эквивалентно аналогичному утверждению для вероятностей q_{ij} (см. выше), однако мы приведем независимое доказательство. Проводя элементарные преобразования

$$\mathcal{P}^*(\mathcal{P}(s)) = \sum_{i=1}^{\infty} p_i^* \mathcal{P}(s)^i = \sum_{i=1}^{\infty} p_i^* \sum_{j=0}^{\infty} p_{ij} s^j = \sum_{j=0}^{\infty} s^j \sum_{i=1}^{\infty} p_i^* p_{ij}$$

и сравнивая члены при s^j, $j=1, 2, \ldots$, в уравнении

$$1 - \mathcal{P}^*(\mathcal{P}(s)) = \mu(1 - \mathcal{P}^*(s)),$$

получаем

$$\sum_{i=1}^{\infty} p_i^* p_{ij} = \mu p_j^*, \quad j=1, 2, \ldots,$$

что и требовалось.

Введем марковскую цепь Z_l, $l=0, 1, 2, \ldots$, с множеством состояний $\mathcal{I}$, переходными вероятностями $\hat{q}_{ij}$ и начальным распределением p_i^*. Нетрудно видеть, что равномерно по $i_0, i_1, \ldots, i_{k-1} \in \mathcal{I}$, $k \in [0, n-k_n]$, $k_n \to \infty$ выполняются соотношения

$$p_{Z_0^{(n)}, \ldots, Z_k^{(n)}}(i_0, \ldots, i_k) \sim p_{Z_0, \ldots, Z_k}(i_0, \ldots, i_k).$$

Предельное поведение марковской цепи Z_l, $l=0, 1, 2, \ldots$, рассматривается в 14.28, оно такое же, как у цепи Y_l, $l=1, 2, \ldots$.

14.27. Задача. Рассмотрим ветвящийся процесс $X_0(\omega) = i > 1$, $X_1(\omega)$, $X_2(\omega)$, $\ldots$ с п. ф. $\mathcal{P}(s)$, $\mu = \mathcal{P}'(1-) < 1$, $\mathcal{P}''(1-) < \infty$. Показать, что

$$\mathbf{P}(\{\omega : X_n(\omega) = j\} \mid \Lambda_n) \to p_j^*, \quad j=1, 2, \ldots,$$

где $\Lambda_n = \{\omega : X_n(\omega) > 0\}$, распределение вероятностей p_j^*, $j=1, 2, \ldots$, то же, что и в случае $i=1$ (см. 14.26).

[П. ф. сл. в. $(X_n|\Lambda_n)$ равна

$$(1-p_{i0}^{(n)})^{-1}\sum_{\{j>0\}}p_{ij}^{(n)}s^j=(1-p_{i0}^{(n)})^{-1}(\mathcal{P}_n(s)^i-p_{i0}^{(n)})=$$

$$=1-(1-p_{i0}^{(n)})^{-1}(1-\mathcal{P}_n(s)^i)=1-(1-(p_{10}^{(n)})^i)(1-\mathcal{P}_n(s)^i)=$$

$$=1-\frac{1-\mathcal{P}_n(s)}{1-p_{10}^{(n)}}\cdot\frac{1+\mathcal{P}_n(s)+\ldots+\mathcal{P}_n(s)^{i-1}}{1+p_{10}^{(n)}+\ldots+(p_{10}^{(n)})^{i-1}}\to 1-(1-\mathcal{P}^*(s)).]$$

14.28. Задача. Показать, что для марковской цепи с переходными вероятностями $\hat{q}_{ij}$, определенными в 14.26, выполняются соотношения

$$\widehat{q}_{ij}^{(k)}\to q_j,\quad k\to\infty,\ i,\ j\in\mathcal{J},$$

где q_j, $\mathcal{J}$ введены в 14.26.

[Прежде всего заметим, что из уравнений для стационарных вероятностей q_j марковской цепи Y_l, $l=0,\ 1,\ 2,\ \ldots$, из 14.26

$$q_j=\sum_{\{i\}}q_ip_{ij},\quad j=1,\ 2,\ \ldots,$$

вытекает, что $q_j>0$, если при некотором i имеем $q_i>0$, $p_{ij}>0$. Так как $q_i=p_i^*\theta_i$, то

$$p_{i_k}^*p_{i_ki_{k-1}}\ldots p_{i_1i_0}>0\Rightarrow p_{i_{k-1}}^*>0,\ \ldots,\ p_{i_0}^*>0$$

и, следовательно, $i_k,\ i_{k-1},\ \ldots,\ i_0\in\mathcal{J}$. Поэтому

$$\widehat{q}_{ij}^{(k)}=-\sum_{\{i_1,\ldots,i_{k-1}\in\mathcal{J}\}}\widehat{q}_{ii_1}\widehat{q}_{i_1i_2}\ldots\widehat{q}_{i_{k-1}j}=$$

$$=p_j^*(p_i^*)^{-1}\sum_{\{i_1,\ldots,i_{k-1}\in\mathcal{J}\}}p_{ji_{k-1}}p_{i_{k-1}i_{k-2}}\ldots p_{i_1i}\mu^{-k}=$$

$$=p_j^*(p_i^*)^{-1}\sum_{\{i_1,\ldots,i_{k-1}>0\}}p_{ji_{k-1}}\ldots p_{i_1i}\mu^{-k}=p_j^*(p_i^*)^{-1}p_{ji}^{(k)}\mu^{-k}.$$

Из 14.27 имеем

$$p_{ji}^{(k)}\mu^{-k}=(p_{ji}^{(k)}/(1-p_{j0}^{(k)}))(1-p_{j0}^{(k)})\mu^{-k}\to p_i^*\theta j,\quad k\to\infty,$$

откуда и следует требуемый результат.]

14.29. Задача. Исследовать предельное поведение распределений начального ($l\leqslant k$) и конечного ($l\geqslant n-k$) отрезков условного ветвящегося процесса $V_l^{(n)}=(X_l|\Lambda_n^0)$, $l=0,\ 1,\ \ldots,n$, где X_l, $l=0,\ 1,\ 2,\ \ldots$, — ветвящийся процесс с $\mu<1$, $\mathcal{P}''(1-)<\infty$,

$$\Lambda^0{}_n=\{\omega:X_n(\omega)>0,\ X_{n+1}(\omega)=0\}.$$

Полагая для определенности $X_0=1$, имеем при $n\to\infty$ и фиксированных k, $i_1, \ldots, i_k>0$ с учетом $1-p_{i0}^{(n)}\sim\theta\mu^n i$

$$p_{V_1^{(n)}\ldots V_k^{(n)}}(i_1, \ldots, i_k)=p_{1i_1}\ldots p_{i_{k-1}i_k}\cdot(p_{i_k0}^{(n-k+1)}-p_{i_k0}^{(n-k)})\times$$

$$\times(p_{10}^{(n+1)}-p_{10}^{(n)})^{-1}\to p_{1i_1}\ldots p_{i_{k-1}i_k}\mu^{-k}i_k=p_{Y_1,\ldots,Y_k}(i_1, \ldots, i_k),-$$

где марковская цепь Y_l, $l=0, 1, 2, \ldots$, та же, что и в 14.26.

Аналогично при $i_0, i_1, \ldots, i_k>0$ получаем

$$p_{V_{n-k}^{(n)},\ldots,V_n^{(n)}}(i_k, \ldots, i_0)=p_{1i_k}^{(n-k)}p_{i_ki_{k-1}}\ldots p_{i_1i_0}\cdot p_{i_00}\times$$

$$\times(p_{10}^{(n+1)}-p_{10}^{(n)})^{-1}\to p_{i_k}^* p_{i_ki_{k-1}}\ldots p_{i_1i_0}\mu^{-k}p_{i_00}(1-\mu)^{-1}=$$

$$=p_{i_0}^* p_{i_00}(1-\mu)^{-1}\hat{q}_{i_0i_1}\hat{q}_{i_1i_2}\ldots\hat{q}_{i_{k-1}i_k}.$$

Сравнивая в уравнении $1-\mathscr{P}^*(\mathscr{P}(s))=\mu(1-\mathscr{P}^*(s))$ члены, не содержащие s (см. конец примера 14.26), получаем

$$\sum_{\{i>0\}} p_i^* p_{i0}=1-\mu,$$

так что предельное распределение сл. в. $V_n^{(n)}, \ldots, V_{n-k}^{(n)}$ совпадает с распределением отрезка марковской цепи с начальным распределением $q_i^*=p_i^*p_{i0}(1-\mu)^{-1}$, $i=1, 2, \ldots$, и переходными вероятностями $\hat{q}_{ij}$, определенными в 14.26.]

14.30. Задача. Пусть $X_0(\omega)=1$, $X_1(\omega)$, $X_2(\omega)$, $\ldots$ — ветвящийся процесс с п. ф. $\mathscr{P}(s)$, $\mu=\mathscr{P}'(1-)>1$, $\mathscr{P}''(1-)<\infty$. Положим $\xi_n(\omega)=X_n(\omega)\mu^{-n}$. Вычислить $\mathbf{M}(\xi_{n+1}-\xi_n)^2$, воспользовавшись 9.24.

[Воспользовавшись тем, что $\mathbf{M}(\xi_{n+1}|\xi_n)=\xi_n$, получаем

$$\mathbf{M}\xi_{n+1}\xi_n=\mathbf{M}\mathbf{M}(\xi_{n+1}\xi_n|\xi_n)=\mathbf{M}(\xi_n\mathbf{M}(\xi_{n+1}|\xi_n))=\mathbf{M}\xi_n^2,$$

$$\mathbf{M}(\xi_{n+1}-\xi_n)^2=\mathbf{M}\xi_{n+1}^2-2\mathbf{M}\xi_{n+1}\xi_n+\mathbf{M}\xi_n^2=\mathbf{M}\xi_{n+1}^2-\mathbf{M}\xi_n^2.$$

Из 14.11 имеем

$$\mathbf{M}\xi_n^2=\mu^{-2n}(\mathbf{M}X_n(X_n-1)+\mathbf{M}X_n)=(\mu(\mu-1))^{-1}\mathscr{P}''(1-)(1-\mu^{-n})+$$

$$+\mu^{-n}=(\mu(\mu-1))^{-1}\mathscr{P}''(1-)-((\mu(\mu-1))^{-1}\mathscr{P}''(1-)-1)\mu^{-n},$$

$$\mathbf{M}\xi_{n+1}^2-\mathbf{M}\xi_n^2=((\mu(\mu-1))^{-1}\mathscr{P}''(1-)-1)(1-\mu^{-1})\mu^{-n}.]$$

14.31. Задача. В условии 14.30 показать, что при $\delta>\mu^{-1/2}$ с вероятностью 1 происходит лишь конечное число событий $\{\omega: |\xi_{n+1}(\omega)-\xi_n(\omega)|>\delta^n\}$. Вывести отсюда, что последовательность сл. в. $\xi_n(\omega)$, $n=1, 2, \ldots$, сходится с вероятностью 1 к некоторой сл. в. $\xi(\omega)$.

[По неравенству Чебышева при $1>\delta>\mu^{-1/2}$

$$\mathbf{P}(\omega: |\xi_{n+1}(\omega)-\xi_n(\omega)|>\delta^n) \leqslant \mathrm{const}\cdot(\delta\mu^{1/2})^{-n}.$$

Применяя лемму Бореля—Кантелли (см. 12.11), получаем первое из утверждений 14.31. Если ω таково, что при всех $n>N$

$$|\xi_{n+1}(\omega)-\xi_n(\omega)| \leqslant c\cdot\delta^n,$$

то, очевидно, последовательность $\xi_n(\omega)$ фундаментельна и, следовательно, существует

$$\lim_{n\to\infty}\xi_n(\omega)=\xi(\omega).$$

Но, как мы установили, такие ω составляют подмножество из Ω меры 1.]

14.32.* Задача. В условии 14.30 показать, что

$$\mathbf{M}(\xi_{n+k}-\xi_n)^2=\frac{\mathcal{P}''(1-)+\mu-\mu^2}{\mu(\mu-1)}(1-\mu^{-k})\mu^{-n}, \quad k=1,\ 2,\ \ldots .$$

Отсюда и из 14.31 вывести с помощью леммы Фату (10.22), что $M\xi^2<\infty$, а при $n\to\infty$ выполнены соотношения

$$\mathbf{M}(\xi-\xi_n)^2\to 0, \quad \xi_n \xrightarrow{\mathbf{P}} \xi, \quad \mathbf{M}|\xi-\xi_n|\to 0.$$

Показать, что $\mathbf{M}\xi=1$ (Сл. в. ξ определена в 14.31).

$$[\mathbf{M}(X_{n+k}|\{\omega: X_n(\omega)=i\})=\sum_{\{j\}} jp_{ij}^{(k)}=\frac{d}{ds}\mathcal{P}_k(s)^i|_{s=1-}=i\mu^k,$$

так что $\mathbf{M}(\xi_{n+k}|\xi_n)=\xi_n$, и первое утверждение в 14.32 устанавливается аналогично 14.30. Далее (см. 14.31, 10.22, 14.11)

$$\mathbf{M}\xi^2=\mathbf{M}(\lim_{n\to\infty}\xi_n^2(\omega))\leqslant \varliminf_{n\to\infty}\mathbf{M}\xi_n^2=\mathcal{P}''(1-)(\mu^2-\mu)^{-1},$$

$$\mathbf{M}(\xi-\xi_n)^2=\mathbf{M}(\varliminf_{k\to\infty}(\xi_{n+k}-\xi_n)^2)\leqslant\varliminf_{k\to\infty}\mathbf{M}(\xi_{n+k}-\xi_n)^2=$$

$$=(\mathcal{P}''(1-)+\mu-\mu^2)(\mu(\mu-1))^{-1}\mu^{-n}.$$

В случае $\mathbf{M}(\xi-\xi_n)^2\to 0$ при $n\to\infty$ говорят, что последовательность сл. в. ξ_n *сходится в среднеквадратичном* к сл. в. ξ. Из неравенства $\mathbf{M}|X|\leqslant(\mathbf{M}X^2)^{1/2}$ (см. 13.46) заключаем, что сходимость в среднеквадратичном влечет за собой *сходимость в среднем*:

$$\mathbf{M}|\xi-\xi_n|\to 0, \quad n\to\infty.$$

Поскольку $|\mathbf{M}\xi-\mathbf{M}\xi_n|\leqslant\mathbf{M}|\xi-\xi_n|$, $\mathbf{M}\xi_n=1$, то $\mathbf{M}\xi=1$.]

14.33. Задача. Обозначим $\psi(\lambda)$ преобразование Лапласа сл. в. ξ из 14.31. Показать, что $\psi(\lambda)$ удовлетворяет уравнению

$$\psi(\mu\lambda)=\mathcal{P}(\psi(\lambda)).$$

Вывести отсюда, что $\mathbf{P}(\omega:\xi(\omega)=0)=p$, где p — меньший корень уравнения $\mathscr{P}(s)=s$, $s\in[0, 1]$.

[Положим

$$\psi_n(\lambda)=\mathbf{M}e^{-\lambda\xi_n}=\mathbf{M}\exp(-\lambda X_n\mu^{-n})=\mathscr{P}_n(\exp(-\lambda\mu^{-n})).$$

Из основного функционального уравнения

$$\mathscr{P}_{n+1}(s)=\mathscr{P}(\mathscr{P}_n(s))$$

получаем для преобразований Лапласа соотношение

$$\psi_{n+1}(\mu\lambda)=\mathscr{P}(\psi_n(\lambda)).$$

Поскольку $\xi_n\xrightarrow{\mathrm{P}}\xi$ (см. 14.32), то $\xi_n\xrightarrow{\mathscr{D}}\xi$ при $n\to\infty$ (см. 13.4) и, следовательно, $\psi_n(\lambda)\to\psi(\lambda)$ (см. 13.27). Переходя к пределу при $n\to\infty$, получаем

$$\psi(\mu\lambda)=\mathscr{P}(\psi(\lambda)).$$

Далее заметим, что

$$\psi(\infty)=\lim_{\lambda\to\infty}\psi(\lambda)=\lim_{\lambda\to\infty}\int_0^\infty e^{-\lambda x}\,dF(x)=F(0)-F(0-).$$

Но $\psi(\infty)=\mathscr{P}(\psi(\infty))$, а поскольку $\mathbf{M}\xi=1$ (см. 14.32), то $\psi(\infty)=$ $=F(0)-F(0-)<1$, т. е. $\psi(\infty)$ есть меньший корень уравнения $\mathscr{P}(s)=s$.]

14.34. Пример (ветвящийся процесс, $\mu>1$). Рассмотрим предельное поведение ветвящегося процесса с $\mu=\mathscr{P}'(1-)>1$, $\mathscr{P}''(1-)<\infty$. Как мы видели в 14.10, с вероятностью p, являющейся меньшим корнем уравнения $\mathscr{P}(s)=s$, $0\leqslant s<1$, процесс вырождается, т. е. при некотором n (зависящем от случая) $X_n(\omega)=$ $=0$. Условный процесс $Y_k=(X_k|\Lambda)$, $k=0, 1, 2, \ldots$, $\Lambda=\{\omega:X_n(\omega)=$ $=0$ при некотором $n\}$ имеет конечномерные распределения (при $Y_0=1$)

$$p_{Y_1,\ldots,Y_k}(i_1, \ldots, i_k)=p_{1i_1}p_{i_1i_2}\ldots p_{i_{k-1}i_k}\min\{p^{i_k-1}, 1\}$$

и представляет собой однородную марковскую цепь с переходными вероятностями

$$q_{i0}=p_{i0},\ i\geqslant0,\ q_{ij}=p_{ij}p^{j-i},\quad i, j>0,$$

где p_{ij} — переходные вероятности процесса X_k, $k=0, 1, 2, \ldots$, $X_0=1$, p — вероятность его вырождения. Состояние 0 — поглощающее для марковской цепи Y_k, $k=0, 1, 2, \ldots$. Вероятность того, что процесс, выходящий из 1, не поглощается к моменту времени n, равна

$$\sum_{\{i_1,\ldots,i_n>0\}}p_{1i_1}\ldots p_{i_{n-1}i_n}p^{i_n-1}=\sum_{i=1}^\infty p_{1i}^{(n)}p^{i-1}=$$

$$=p^{-1}(\mathscr{P}_n(p)-\mathscr{P}_n(0))=p^{-1}(p-\mathscr{P}_n(0)).$$

Учитывая, что $\mathcal{P}_n(0)\to p$ при $n\to\infty$ и

$$p-\mathcal{P}_n(0)=\mathcal{P}(p)-\mathcal{P}(\mathcal{P}_{n-1}(0))=$$

$$=\mathcal{P}'(p)(p-\mathcal{P}_{n-1}(0))+\frac{1}{2}\mathcal{P}'(\theta)(p-\mathcal{P}_{n-1}(0))^2,$$

выводим, как и в 14.15, что

$$p-\mathcal{P}_n(0)\sim c_1\mathcal{P}'(p)^n,\quad n\to\infty,\ 0<\mathcal{P}'(p)<1.$$

Таким образом, для сл. в. T_1 — времени до поглощения марковской цепи $Y_0=1$, Y_1, Y_2, ... —

$$\mathbf{P}(\omega:T_1(\omega)>n)\sim c_1\mathcal{P}'(p)^n,\quad n\to\infty,$$

так что поглощение происходит с *экспоненциальной скоростью*; в частности, конечны $\mathbf{M}T_1$ и $\mathbf{D}T_1$. Для процесса, начинающегося из состояния k, получаем точно так же, что

$$\mathbf{P}(\omega:T_k(\omega)>n)\sim c_k\mathcal{P}'(p)^n,\quad n\to\infty.$$

На множестве $\bar{\Lambda}=\{\omega:X_n(\omega)>0,\ n=1,\ 2,\ \ldots\}$ поведение ветвящегося процесса X_n, $n=0,\ 1,\ 2,\ \ldots$, определяется в 14.31. Как там показано, с вероятностью 1

$$\xi_n(\omega)=X_n(\omega)\cdot\mu^{-n}\to\xi(\omega),\quad n\to\infty,$$

где сл. в. $\xi(\omega)$ равняется нулю на некотором множестве A меры $\mathbf{P}(A)=p$, а на его дополнении $\bar{A}$ $\xi(\omega)>0$. Очевидно, что множество A отличается от Λ не более чем на подмножество меры нуль. Следовательно, если пренебречь некоторым подмножеством меры нуль, то траектории процесса $X_n(\omega)$, $n=0,\ 1,\ 2,\ \ldots$, разбиваются на два класса: одни поглощаются за конечное время в нуле, для других выполняется соотношение $X_n(\omega)\sim\xi(\omega)\mu^n$, $\xi(\omega)>0$; эти возможности имеют вероятности соответственно p и $1-p$.

14.35. Пример (ветвящийся процесс с иммиграцией). Рассмотрим следующее обобщение процесса Гальтона—Ватсона. Представим себе, что к обычному ветвящемуся процессу в моменты времени $n=1,\ 2,\ \ldots$ добавляется случайное число Y_n частиц, которые принимают участие в дальнейшем процессе размножения по тем же самым законам. Обозначим Z_n число частиц в момент времени n в описанном *процессе с иммиграцией*, предполагая, что сл. в. Y_i, $i=1,\ 2,\ \ldots$, независимы и одинаково распределены с п. ф. $\mathcal{Q}(s)$. Сл. в. Z_n представима в виде суммы

$$Z_n=X_n+V_{n-1}^{(1)}+V_{n-2}^{(2)}+\ldots+V_1^{(n-1)}+Y_n,\quad n=1,\ 2,\ \ldots,\ Z_0=1,$$

где X_n — число потомков начальной частицы в n-й момент времени в обычном процессе Гальтона—Ватсона, $V_{n-k}^{(k)}$— число потомков в n-й момент времени от Y_k частиц, иммигрировавших в момент k, $k=1,\ 2,\ \ldots,\ n-1$. Случайные величины в сумме, составляющей Z_n, по определению процесса являются независимыми, так что п. ф. сл. в. Z_n равняется произведению п. ф. слагаемых. Но

$$\mathbf{M}s^{V_m^{(k)}}=\mathbf{M}\mathbf{M}(s^{V_m^{(k)}}\,|Y_k)=\mathbf{M}(\mathcal{P}_m(s))^{Y_k}=Q(\mathcal{P}_m(s)),$$

так что

$$\mathbf{M}s^{Z_n}=\mathcal{P}_n(s)\prod_{k=1}^{n}Q(\mathcal{P}_{n-k}(s)),\quad \mathcal{P}_0(s)=s,$$

где $\mathcal{P}_n(s)=\mathcal{P}(\mathcal{P}(\ldots\mathcal{P}(s)\ldots))$ n раз, $\mathcal{P}(s)$ — п. ф. числа непосредственных потомков одной частицы. Нетрудно формализовать приведенное определение процесса Z_n, $n=0, 1, 2, \ldots$, так, как это было сделано в 9.24 в случае процесса Гальтона — Ватсона.

Другой способ определения марковской цепи Z_n, $n=0, 1, 2, \ldots$, состоит в указании ее переходных вероятностей

$$p_{Z_{n+1}|Z_n}(j\,|\,i)=\pi_{ij},\quad \sum_{j=0}^{\infty}\pi_{ij}s^j=\mathcal{P}(s)^i\,Q(s).$$

Полагая $Z_0=1$,

$$\mathcal{R}_n(s)=\mathbf{M}s^{Z_n}=\sum_{j=0}^{\infty}\pi_{1j}^{(n)}s^j,$$

имеем

$$\mathcal{R}_{n+1}(s)=\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\pi_{1i}^{(n)}\pi_{ij}s^j=\sum_{i=0}^{\infty}\pi_{1i}^{(n)}\sum_{j=0}^{\infty}\pi_{ij}s^j=\mathcal{R}_n(\mathcal{P}(s))\,Q(s),$$

что приводит к той же самой п. ф. для сл. в. Z_n, что и выше.

Далее предположим, что $Q(0)=\mathbf{P}(\omega:Y_1(\omega)=0)<1$, т. е. иммиграция в самом деле имеет место, а также, что $Q'(1-)<\infty$, и рассмотрим случай $\mu=\mathcal{P}'(1-)=1$, $\mathcal{P}'''(1-)<\infty$. Как мы знаем, процесс без иммиграции в этих предположениях с вероятностью 1 когда-то попадает в состояние 0 (и в нем остается). Рассмотрим вопрос о попадании (возвращении) в какое-либо состояние в процессе с иммиграцией. Положим

$$f_{ij}^{(n)}=\sum_{\{i_1\neq j,\, i_2\neq j,\ldots,\, i_{n-1}\neq j\}}\pi_{ii_1}\pi_{i_1i_2}\ldots\pi_{i_{n-1}j},$$

т. е. $f_{ij}^{(n)}$ есть вероятность того, что марковская цепь, выходящая из состояния i на n-м шаге, впервые попадает в состояние j. Отсюда имеем для произвольной марковской цепи

$$\pi_{ij}^{(n)}=\sum_{k=1}^{n}\sum_{\{i_1\neq j,\ldots,i_{k-1}\neq j,\, i_k=j,\, i_{k+1},\ldots,i_n\}}\pi_{1i_1}\ldots\pi_{i_{n-1}j}=$$

$$=f_{ij}^{(1)}\pi_{jj}^{(n-1)}+f_{ij}^{(2)}\pi_{jj}^{(n-2)}+\ldots+f_{ij}^{(n-1)}\pi_{jj}+f_{ij}^{(n)}$$

(событие $\{\omega : Z_n(\omega)=j\}$ разбито на непересекающиеся части в соответствии с тем, когда произошло первое попадание в состояние j). Домножая обе части полученного уравнения на s^n и суммируя по n от 1 до ∞, получаем

$$\mathcal{P}_{ij}(s)=\sum_{n=1}^{\infty} \pi_{ij}^{(n)} s^n=\mathcal{F}_{ij}(s)+\mathcal{F}_{ij}(s)\mathcal{P}_{jj}(s),$$

$$\mathcal{F}_{ij}(s)=\sum_{n=1}^{\infty} f_{ij}^{(n)} s^n.$$

При $i=j$ имеем

$$\mathcal{P}_{jj}(s)=\mathcal{F}_{jj}(s)/(1-\mathcal{F}_{jj}(s)).$$

Отсюда видно, что $\mathcal{F}_{ii}(1-)=1$ тогда и только тогда, когда $\mathcal{P}_{ij}(1-)=\infty$, т. е. возвращение в состояние j происходит с вероятностью 1, только если расходится ряд

$$\sum_{n=1}^{\infty} \pi_{jj}^{(n)}=\infty.$$

Если $\mathcal{F}_{ii}(s)\not\equiv 0$, т. е. из состояния i можно с положительной вероятностью попасть в состояние j на некотором шаге, то $\mathcal{P}_{jj}(1-)=\infty$ и тогда и только тогда, когда $\mathcal{P}_{ij}(1-)=\infty$, так что возвращение в состояние j с вероятностью 1 происходит, только если

$$\sum_{n=1}^{\infty} \pi_{ij}^{(n)}=\infty.$$

Как показывают рассуждения на с. 318, если $\mathcal{P}_{ij}(1-)=\infty$ и $\mathcal{F}_{ji}(1-)>0$, то $\mathcal{F}_{ij}(1-)=1$.

Возвращаясь к ветвящемуся процессу с иммиграцией, предположим вначале, что $Q(0)>0$, т. е. с положительной вероятностью иммигранты в данный момент не поступают. В этом случае, очевидно, $\pi_{i0}^{(n)}>0$ при любых n, i. Найдем асимптотику вероятностей

$$\pi_{10}^{(n)}=\mathcal{P}_n(0)\prod_{k=1}^{n} Q(\mathcal{P}_{n-k}(0)), \quad n\to\infty, \ \mathcal{P}_0(s)=s.$$

Логарифмируя и используя асимптотику 14.17, получаем

$$\ln \pi_{10}^{(n)}=\ln \mathcal{P}_n(0)+\sum_{k=1}^{n} \ln Q(\mathcal{P}_{n-k}(0))= -\sum_{k=1}^{n} (2Q'(1-)/\mathcal{F}''(1-))\,k^{-1}+$$
$$+o(\ln n)=-(2Q'(1-)/\mathcal{F}''(1-))\ln n+o(\ln n).$$

Следовательно, ряд

$$\sum_{n=1}^{\infty} \pi_{10}^{(n)}$$

расходится, если $2Q'(1-)/\mathcal{P}''(1-)<1$, и сходится при $2Q'(1-)/\mathcal{P}''(1-)>1$. При $\mathcal{F}_{01}(s)\not\equiv 0$ вопрос о попадании с вероятностью 1 в состояние 0 определяется значением $Q'(1-)$ математического ожидания числа иммигрантов, за исключением пограничного случая $Q'(1-)=\mathcal{P}''(1-)/2$. При дополнительном предположении $Q''(1-)<\infty$ разложение по формуле Тейлора в предыдущих выкладках до второго члена показывает, что в случае $Q'(1-)=\mathcal{P}''(1-)/2$ имеет место расходимость ряда из $\pi_{10}^{(n)}$ $n=1, 2, \ldots$, и, следовательно, цепь попадает в состояние 0 с вероятностью 1. Поскольку

$$\pi_{i0}^{(n)}=\pi_{10}^{(n)}\cdot(\mathcal{P}_n(0))^{i-1}, \quad i\geqslant 0,$$

то все эти выводы сохраняются при любом начальном состоянии (или начальном распределении).

Отметим, что процесс с иммиграцией не может попасть в состояние 0 ранее, чем выродился процесс размножения начальных частиц. Следовательно, обозначив $T_i(\omega)$ время до первого попадания в состояние 0 процесса, начавшегося с i частиц, получаем при $n\to\infty$

$$\mathbf{P}(\omega : T_i(\omega)>n)\geqslant 1-\mathcal{P}_n(0)^i\sim 2i/(\mathcal{P}''(1-)n),$$

так что $\mathbf{M}T_1=\infty$.

Сняв предположение $Q(0)>0$, обратимся к марковской цепи (ср. выше определение Z_n)

$$\widetilde{Z}_n=X_n+V_{n-1}^{(1)}+V_{n-2}^{(2)}+\ldots+V_1^{(n-1)}, \quad n=1, 2, \ldots,$$

так что Z_n представляет собой сумму $\widetilde{Z}_n$ и иммигрантов, прибывших в момент n. К п. ф.

$$\mathbf{M}s^{\widetilde{Z}_n}=\mathcal{P}_n(s)^i\prod_{k=1}^{n-1} Q(\mathcal{P}_{n-k}(s)), \quad \widetilde{Z}_0=i,$$

применимы предыдущие построения. Возвращение марковской цепи $\widetilde{Z}_n$, $n=0, 1, 2, \ldots$, в состояние 0 происходит с вероятностью 1, если $2Q'(1-)/\mathcal{P}''(1-)<1$ (или $=1$, но $Q''(1-)<\infty$). При этом ряд из $\widetilde{\pi}_{i0}^{(n)}$ расходится и, воспользовавшись неравенством

$$\pi_{ij}^{(n)}\geqslant\widetilde{\pi}_{i0}^{(n)}q_j, \quad Q(s)=\sum_{k=1}^{\infty} q_k s^k, \quad \widetilde{\pi}_{i0}^{(n)}=p_{\widetilde{Z}_n|\widetilde{Z}_0}(0|i),$$

получаем, что при $q_j>0$ расходится ряд из $\pi_{ij}^{(n)}$, $n=1, 2, \ldots$, и, следовательно, возвращение в состояние j происходит с вероятностью 1.

Обозначим

$$f_{ij}=\sum_{n=1}^{\infty} f_{ij}^{(n)}=\mathcal{F}_{ij}(1)$$

вероятность, выходя из состояния i, когда-то попасть в состояние j. Допустим, что $f_{ij}=f_{jj}=1$. Тогда попадание в состояние j (начиная из i) происходит бесконечно часто с вероятностью 1. В самом деле, пусть τ_1 — момент первого попадания в j, Δ_k — интервал между k-м и $(k+1)$-м попаданиями в j. Очевидно, что

$$\mathbf{P}(\omega : \tau_1(\omega)=n_1,\ \Delta_k(\omega)=n_k,\ 2\leqslant k\leqslant m)=f_{ij}^{(n_1)}f_{jj}^{(n_2)}\ \ldots\ f_{jj}^{(n_m)},$$

т. е. сл. в. $\tau_1, \Delta_2, \Delta_3, \ldots$ независимы, а сл. в. $\Delta_2, \Delta_3, \ldots$ одинаково распределены. Поэтому

$$\mathbf{P}(\omega : \Delta_k(\omega)<\infty,\ k=2, 3, \ldots)=\lim_{n\to\infty}\mathbf{P}(\omega : \Delta_k(\omega)<\infty,\ k\leqslant n)=1.$$

Также нетрудно видеть, что отрезки марковской цепи $Z_i(\omega)$, $\tau_{k-1}\leqslant i\leqslant\tau_k$, $\tau_k=\tau_{k-1}+\Delta_k$, $k=1, 2, \ldots, \tau_0=0$, $\Delta_1=\tau_1$, независимы, а начиная с $k=2$, и одинаково распределены. Таким образом, траекторию марковской цепи можно разбить на независимые циклы между последовательными попаданиями в состояние j. Подчеркнем, что эти выводы относятся к произвольной марковской цепи, для которой $f_{ij}=f_{jj}=1$.

Из сказанного можно получить многие интересные выводы. Так, в предположении $f_{jj}=1$ допустим, что $f_{jl}>0$, т. е. из состояния j с положительной вероятностью можно когда-то попасть в состояние l. События A_k, что попадание в состояние l произойдет между k-м и $(k+1)$-м попаданиями в состояние j, $k=1, 2, \ldots$, независимы и имеют одну и ту же положительную вероятность. Поэтому вероятность того, что произойдет хотя бы одно из событий $A_1, A_2, \ldots$ (а значит и бесконечно много), равна 1, т. е. $f_{jl}=1$. Точно так же получаем, что $f_{lj}=1$.

Обратимся теперь к случаю $\mu=\mathcal{P}'(1-)<1$, $\mathcal{P}''(1-)<\infty$, $\mathcal{Q}(0)>0$. Покажем при дополнительном ограничении $\mathcal{Q}''(1-)<<\infty$, что

$$\mathcal{R}_n(s)=\mathbf{M}s^{Z_n}\to\mathcal{R}(s),\quad 0\leqslant s\leqslant 1,\quad n\to\infty,$$

где $\mathcal{R}(s)$ — некоторая вероятностная производящая функция. Таким образом, марковская цепь $Z_0=1, Z_1, Z_2, \ldots$ при $n\to\infty$ входит в стационарный режим. Запишем

$$\ln\mathcal{R}(s)=\ln\mathcal{P}_n(s)+\sum_{k=1}^{n}\ln\mathcal{Q}(\mathcal{P}_{n-k}(s)),\quad \mathcal{P}_0(s)=s.$$

Воспользуемся соотношением

$$1-\mathcal{P}_n(s)=\theta\mu^n+O(1)\mu^{2n},\quad n\to\infty,$$

выполняющимся равномерно по $s \in [0, 1]$ (ср. 14.15, 14.21), и разложением (также действующим равномерно по s)

$$\ln Q(\mathcal{P}_m(s)) = Q'(\theta_m)(1 - \mathcal{P}_m(s)) + O(1)(1 - \mathcal{P}_m(s))^2, \quad m \to \infty.$$

Таким образом,

$$\ln \mathcal{R}_n(s) = \sum_{k=0}^{n-1} \ln Q(\mathcal{P}_k(s)) + \ln \mathcal{P}_n(s),$$

причем ряд из членов $\ln Q(\mathcal{P}_k(s))$ равномерно сходится. Следовательно, функция $\mathcal{R}(s) = \lim\limits_{n\to\infty} \mathcal{R}_n(s)$ непрерывна. Из 13.29 заключаем, что $\mathcal{R}(s)$ — вероятностная п. ф.

Если процесс начался с i частиц, то, очевидно, имеет место тот же самый предельный результат. Положив

$$\mathcal{R}(s) = \sum_{j=0}^{\infty} \pi_j s^j,$$

можно записать полученное утверждение в виде

$$\pi_{ij}^{(n)} \to \pi_j, \quad j = 0, 1, 2, \ldots, \quad n \to \infty.$$

Отметим, что для j, таких, что $\pi_j > 0$, имеем

$$\sum_{n=1}^{\infty} \pi_{ij}^{(n)} = \infty, \quad \sum_{n=1}^{\infty} \pi_{jj}^{(n)} = \infty,$$

так что с вероятностью 1 эти состояния посещаются бесконечно часто.

14.36.* Пример (простое блуждание и ветвящийся процесс). Рассмотрим простое блуждание S_n, $n=0, 1, 2, \ldots$, $S_0=0$ с вероятностью шага вправо $p \geqslant 1/2$. Введем момент T_r первого достижения положения r, $r=1, 2, \ldots$. Сл. в. T_r, $r=1, 2, \ldots$, конечны с вероятностью 1 (см. 12.15), последовательность S_{T_r+k}, $k=0, 1, 2, \ldots$, представляет собой блуждание начинающееся из точки r, отрезки блуждания $S_{T_r}, S_{T_r+1}, \ldots, S_{T_{r+1}}$, $r=0, 1, 2, \ldots$, $T_0=0$, независимы. Если сдвинуть начало координат в (случайную) точку (T_r, r), то отрезок $S_{T_r}, S_{T_r+1}, \ldots, S_{T_{r+1}}$ будет представлять собой блуждание, выходящее из нуля и остановленное в момент первого достижения единицы (ср. 11.5). С отрезком S_{T_r+m}, $m=0, 1, \ldots, T_{r+1}-T_r$ свяжем последовательность сл. в. $Z_l^{(r)}$, $l = r, r-1, r-2, \ldots$, полагая $Z_l^{(r)}$ равным числу переходов блуждающей частицы из положения l в $l+1$ (на интервале блуждания (T_r, T_{r+1})). Число отрезков путей, таких, что $Z_{r-1}^{(r)} = i_1 > 0, \ldots, Z_{r-k}^{(r)} = i_k > 0$, $Z_{r-k+1}^{(r)} = 0$ было подсчитано в 11.9 и равно

$$C_{i_1+i_2-1}^{i_2} C_{i_2+i_3-1}^{i_3} \ldots C_{i_{k-1}+i_k-1}^{i_k}.$$

Эти пути имеют длину $n=2(i_1+\ldots+i_k)+1$, причем суммарное число шагов блуждающей частицы вверх на единицу больше числа шагов вниз, так что все эти отрезки имеют одинаковую вероятность $p^{(n+1)/2}(1-p)^{(n-1)/2}$. Таким образом, получаем при любых $k\geqslant 0$, $i_1, \ldots, i_k\geqslant 0$:

$$\mathbf{P}(\omega : Z_{r-1}^{(r)}(\omega)=i_1, \ldots , Z_{r-k}^{(r)}(\omega)=i_k, \quad Z_{r-k-1}^{(r)}(\omega)=0)=$$

$$=C_{i_1+i_2-1}^{i_2} \ldots C_{i_{k-1}+i_k-1}^{i_k} p^{(n+1)/2}(1-p)^{(n-1)/2}=$$

$$=p(i_1\,|\,1)\,p(i_2\,|\,i_1) \ldots p(i_k\,|\,i_{k-1})\,p(0\,|\,i_k),$$

$$p(j\,|\,i)=C_{i+j-1}^{j}p^i(1-p)^j, \quad i=1, 2, \ldots , \quad j=0, 1, 2, \ldots .$$

Поскольку $p(j\,|\,i)$, $j=0, 1, 2, \ldots$, представляет собой отрицательное биномиальное распределение, то вероятности $p(j\,|\,i)$, $i>$ >0, $p(0\,|\,0)=1$, $p(j\,|\,0)=0$, $j>1$, образуют систему переходных вероятностей на множестве целых неотрицательных чисел. Отсюда легко выводится (см. 11.9), что сл. в. $Z_r^{(r)}=1$, $Z_{r-1}^{(r)}$, $Z_{r-2}^{(r)}, \ldots$ образуют марковскую цепь с переходными вероятностями $p(j\,|\,i)$ $i, j\geqslant 0$. Поскольку (см. 4.10)

$$\sum_{j=0}^{\infty} p(j\,|\,i)\,s^j=\left(\frac{p}{1-(1-p)\,s}\right)^i = \left(\sum_{j=0}^{\infty} p(j\,|\,1)\,s^j\right)^i, \quad i>0,$$

то последовательность $Z_{r-l}^{(r)}$, $l=0, 1, 2, \ldots$, представляет собой процесс Гальтона—Ватсона **с п. ф.**

$$\mathcal{P}(s)=p/(1-(1-p)s),$$

отвечающей геометрическому распределению $(1-p)^jp$, $j=0, 1, 2, \ldots$, числа непосредственных потомков одной частицы.

Введем последовательность сл. в. $V_k^{(m)}$, $k=m, m-1, \ldots, 1, 0, -1, \ldots$, полагая $V_k^{(m)}$ равным числу переходов блуждающей частицы из положения k в $k+1$ на интервале блуждания от 0 до T_{m+1}, т. е. до первого момента достижения точки $m+1$ (в частности, $V_k^{(0)}=Z_k^{(0)}$). Нетрудно видеть, что

$$V_k^{(m)}=\sum_{r=k}^{m} Z_k^{(r)}, \quad m=0, 1, 2, \ldots ,$$

причем процессы $Z_k^{(r)}$, $k=r, r-1, \ldots$, при различных r независимы, так как определяются независимыми отрезками случайного блуждания S_n, $n=0, 1, 2, \ldots$. Перенумеруем сл. в. $V_k^{(m)}$, $Z_k^{(r)}$ в обратном порядке, полагая при $t=0, 1, \ldots, m$

$$U_l^{(t)}=Z_{m-l}^{(m-t)}, \quad W_l^{(m)}=V_{m-l}^{(m)}=\sum_{t=0}^{\min(l,m)} U_l^{(t)}.$$

Трактуем индекс l как временной параметр. Тогда $U_l^{(t)}$, $l=t$, $t+1, \ldots, 0 \leqslant t \leqslant m$, есть процесс Гальтона—Ватсона с п. ф. $\mathscr{P}(s) = p/(1-(1-p)s)$, начавшийся в момент $l=t$ с одной частицы ($U_t^{(t)} = Z_{m-t}^{(m-t)} = 1$). При различных t, $0 \leqslant t \leqslant m$, процессы $U_l^{(t)}$, $l=t$, $t+1, \ldots$, независимы. Сл. в. $W_l^{(m)}$ представляет собой суммарное число частиц, живущих в момент времени l. Таким образом, последовательность $W_l^{(m)}$, $l=0, 1, 2, \ldots$, представляет собой ветвящийся процесс следующего вида: он начинается с одной частицы, в моменты $l=1, 2, \ldots, m$ происходит иммиграция одной частицы, при $l>m$ иммиграция отсутствует, размножение частиц в течение всего времени происходит по законам процесса Гальтона—Ватсона с п. ф. $\mathscr{P}(s)$, указанной выше.

Неоднородную марковскую цепь $W_l^{(m)}$, $l=0, 1, 2, \ldots$, естественно разбить на однородные участки $l \leqslant m$ и $l \geqslant m$. При $0 \leqslant l \leqslant m$ $W_l^{(m)}$ есть ветвящийся процесс с п. ф. $\mathscr{P}(s)$ и иммиграцией по одной частице для $l \geqslant 1$, его переходные вероятности равны $\pi_{ij} = p(j-1 \mid i)$, $i, j = 1, 2, \ldots$. При $l \geqslant m$ это обычный процесс Гальтона—Ватсона, начавшийся с $W_m^{(m)}$ частиц. По свойству марковской цепи зависимость между двумя участками осуществляется через значение сл. в. $W_m^{(m)}$.

Рассмотрим детальнее отрезок $W_l^{(m)}$, $l \leqslant m$, при $m \to \infty$. Введем W_l, $l=0, 1, 2, \ldots$, — ветвящийся процесс с п. ф. $\mathscr{P}(s)$, иммиграцией по одной частице при $l \geqslant 1$. Если $W_0 = 1$, то процесс $W_l^{(m)}$, $l \leqslant m$, совпадает по распределению с отрезком W_l, $l \leqslant m$. Так как $\mu = \mathscr{P}'(1) = (1-p)/p$, то $\mu = 1$ при $p = 1/2$ и $\mu < 1$ при $p > 1/2$. Применим результаты, полученные в 14.35. П. ф. числа иммигрантов в данном случае равна $\mathscr{Q}(s) = s$. При $p = 1/2$ вычисляем значение

$$2\mathscr{Q}'(1)/\mathscr{F}''(1) = p^2/(1-p)^2 = 1.$$

Поскольку $\mathscr{Q}(0) = 0$, $\mathscr{Q}''(1) = 0$, то выполняются условия $\mathscr{Q}(0) < 1$, $\mathscr{Q}''(1) < \infty$, и процесс W_l, $l=0, 1, 2, \ldots$, из любого состояния с вероятностью 1 когда-либо попадает в состояние 1, так как все состояния марковской цепи W_l сообщающиеся: $f_{ij} > 0$, $i, j = 1, 2, \ldots$, то $f_{ij} = 1$ при любых i, j, т. е. из любого состояния цепь с вероятностью 1 когда-либо попадает в любое другое и возвращается в исходное (см. конец 14.35). Эти же выводы сохраняются при $p > 1/2$. Кроме того, при $p > 1/2$ марковская цепь W_l, $l = 0, 1, 2, \ldots$, сходится к стационарному режиму, причем стационарное распределение строго положительно при всех $j \geqslant 1$.

14.37.* Задача. В условиях 14.36 при $p > 1/2$ положим Y_l равным полному числу переходов блуждания S_n, $n=0, 1, 2, \ldots$, $S_0 = 0$ из положения l в $l+1$, $l=0, 1, 2, \ldots$. Показать, что последовательность сл. в. Y_l, $l=0, 1, 2, \ldots$, образует стационарную марковскую цепь. Показать, что переходные вероятности обращенной цепи $p_{Y_l \mid Y_{l+1}}(j \mid i)$ совпадают с переходными вероятностями $\pi_{ij} =$

$=p(j-1|i)$, определенными в 14.36. Найти распределение вероятностей сл. в. Y_l.

Прежде всего отметим, что, поскольку с вероятностью 1 блуждание уходит на ∞, сл. в. Y_l конечна с вероятностью 1, $l=0, 1, 2, \ldots$. Пусть T_r — момент первого достижения блуждани-

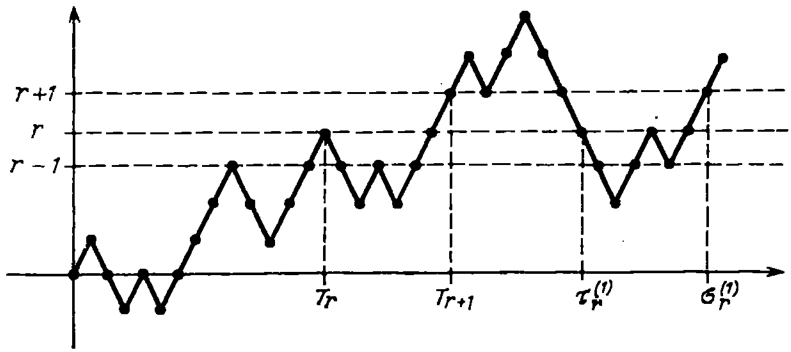

Рис. 36

ем S_n, $n=0, 1, 2, \ldots$, положения $r>0$, $\tau_r^{(1)}$ — момент первого после T_{r+1} попадания в состояние r, $\sigma_r^{(1)}$ — момент первого после T_{r+1} попадания в $r+1$ и далее при $i=2, 3, \ldots$ (рис. 36):

$$\tau_r^{(i)}=\inf\{n : n>\sigma_r^{(i-1)},\quad S_n=r\},$$

$$\sigma_r^{(i)}=\inf\{n : n>\tau_r^{(i)},\quad S_n=r+1\}.$$

Сл. в. $\tau_r^{(i)}$, $\sigma_r^{(i)}$, $i=1, 2, \ldots$, доопределим значением ∞ для тех ω, на которых они не определены предыдущими соотношениями. Обозначим N_r число конечных членов последовательности $\tau_r^{(1)}, \tau_r^{(2)}, \ldots$. Заметим, что если пренебречь событием нулевой вероятности, то из конечности $\tau_r^{(i)}$ следует конечность $\sigma_r^{(i)}$ Нетрудно понять, что отрезки блуждания

$$S_{\tau_r^{(i)}}, S_{\tau_r^{(i)}+1}, \ldots, S_{\sigma_r^{(i)}}, \quad i=1, 2, \ldots, N_r$$

независимы и являются вероятностными копиями друг друга. Формально это означает

$$\mathbf{P}(N_r=l,\ \sigma_r^{(i)}-\tau_r^{(i)}=n_i,\ S_{\tau_r^{(i)}+1}=k_1^{(i)}, \ldots, S_{\sigma_r^{(i)}-1}=k_{n_i-1}^{(i)},\ 1\leqslant i\leqslant l)=$$

$$=\mathbf{P}(N_r=l)\cdot\prod_{i=1}^{l}\mathbf{P}(\sigma_r^{(1)}-\tau_r^{(1)}=n_i,\ S_{\tau_r^{(1)}+1}=k_1^{(i)}, \ldots, S_{\sigma_r^{(1)}-1}=k_{n_i-1}^{(i)}),$$

что непосредственно вытекает из марковского свойства простого случайного блуждания.

С отрезками блужданий на интервалах $(0, T_{r+1})$, $(\tau_r^{(1)}, \sigma_r^{(1)})$, $(\tau_r^{(2)}, \sigma_r^{(2)})$, ... связаны ветвящиеся процессы $Z_l^{(r)}$, $Y_l^{(r,1)}$, $Y_l^{(r,2)}$, ..., $l=r, r-1, \ldots, 0$, где сл. в. $Z_l^{(r)}$ определена в 14.36, $Y_l^{(r,i)}$ определяется по отрезку случайного блуждания на интервале $(\tau_r^{(i)}, \sigma_r^{(i)})$ (при условии конечности этого интервала) точно так же, как сл. в. $Z_l^{(r)}$ определена по отрезку блуждания на $(0, T_{r+1})$. Все эти процессы независимы, $Z_{l-r}^{(r)}$, $l=0, 1, 2, \ldots$, — процесс с иммиграцией, $V_{l-r}^{(r,i)}$, $l=0, 1, 2, \ldots$, — обычный процесс Гальтона—Ватсона с той же п. ф. $\mathscr{P}(s)$, что и $Z_l^{(r)}$. Очевидно,

$$Y_l=Z_l^{(r)}+Y_l^{(r,1)}+\ldots+Y_l^{(r,N_r)}, \quad l=0, 1, \ldots, r,$$

т. е. процесс Y_l, $l=0, 1, \ldots, r$, есть сумма случайного числа N_r+1 независимых ветвящихся процессов. Это означает, что последовательность сл. в. $Y_l^*=Y_{r-l}$, $l=0, 1, \ldots, r$, — ветвящийся процесс с п. ф. $\mathscr{P}(s)=p/(1-(1-p)s)$, начавшийся с N_r+1 частиц, с иммиграцией по одной частице в моменты времени $l=1, \ldots, r$.

Далее заметим, что событие $\{\omega : N_r(\omega)>0\}$ означает, что после момента T_{r+1} первого достижения точки $r+1$ когда-то происходит попадание в точку r. Соответствующую вероятность легко подсчитать, используя рассуждения 12.15, откуда получаем для нее значение $(1-p)/p$. Если событие $\{\omega : N_r(\omega)>0\}$ произошло, то с вероятностью 1 $\sigma_r^{(1)}<\infty$, и блуждание после момента $\sigma_r^{(1)}$ имеет снова возможность попасть в состояние r с вероятностью $(1-p)/p$ и т. д. Следовательно,

$$\mathbf{P}(\omega : N_r(\omega)=n)=\left(\frac{1-p}{p}\right)^n\left(\frac{2p-1}{p}\right), \quad n=0, 1, 2, \ldots .$$

Остается заметить, что $Y_r=N_r+1$ и потому марковская цепь Y_l, $l=0, 1, 2, \ldots$, стационарна и

$$\mathbf{P}(\omega : Y_l(\omega)=i)=\left(\frac{1-p}{p}\right)^{i-1}\frac{2p-1}{p}\equiv\pi_i, \quad i=1, 2, \ldots .$$

Это стационарное распределение марковской цепи с переходными вероятностями

$$\pi_{ij}=p(j-1\,|\,i)=C_{i+j-2}^{j-1}p^i(1-p)^{j-1}, \quad i, j=1, 2, \ldots ;$$

что, впрочем, легко проверяется непосредственно:

$$\sum_{i=1}^{\infty}\pi_i\pi_{ij}=\sum_{i=1}^{\infty}\left(\frac{1-p}{p}\right)^{i-1}\frac{2p-1}{p}C_{i+j-2}^{j-1}p^i(1-p)^{j-1}=$$

$$=\left(\frac{1-p}{p}\right)^{j-1}\frac{2p-1}{p}\sum_{i=1}^{\infty}C_{j+(i-1)-1}^{i-1}p^j(1-p)^{i-1}=\pi_j.$$

Из 14.35, 14.36 имеем, что при $n\to\infty$

$$\sum_{j=1}^{\infty} \pi_{ij}^{(n)} s^j \to \mathcal{P}^*(s) = \sum_{j=1}^{\infty} p_j^* s^j,$$

где $\mathcal{P}^*(s)$ — некоторая вероятностная п. ф. В 14.26 было отмечено, что p_j^*, $j=1, 2, \ldots$, есть стационарное распределение марковской цепи и что оно единственно, т. е. $p_j^* = \pi_j$, $j=1, 2, \ldots$.

Наконец отметим, что

$$p_{Y_{l+1}|Y_l}(i\,|\,j) = \pi_i \pi_{ij}/\pi_j = C_{i+j-2}^{j-1} p^j (1-p)^{i-1} = \pi_{ji},$$

т. е. марковская цепь Y_l, $l=0, 1, 2, \ldots$, и в прямом, и в обратном порядке имеет одни и те же переходные вероятности.]

14.38.* Пример (ветвящийся процесс в случайной среде). Пусть имеется семейство $\mathcal{P}_x(s)$, $x\in\mathcal{X}$, вероятностных п. ф. Возьмем какую-либо последовательность $\mathbf{x} = (x_0, x_1, x_2, \ldots)$ элементов из $\mathcal{X}$ и определим неоднородную марковскую цепь $Z_n^{\mathbf{x}}$, $n=0, 1, 2, \ldots$, с переходными вероятностями

$$\sum_{j=0}^{\infty} p_{Z_{n+1}^{\mathbf{x}}|Z_n^{\mathbf{x}}}(j\,|\,i)\, s^j = \sum_{j=0}^{\infty} p_{ij}(x_n)\, s^j = \mathcal{P}_{x_n}(s)^i, \quad i\geqslant 0.$$

Интерпретируя состояние i как число частиц в процессе размножения, при котором частицы, живущие в момент n, производят непосредственных потомков по законам процесса Гальтона—Ватсона с п. ф. $\mathcal{P}_{x_n}(s)$, называем марковскую цепь $Z_n^{\mathbf{x}}$, $n=0, 1, 2, \ldots$, *неоднородным ветвящимся процессом.* Значение параметра $x=x_n$ можно трактовать как «состояние среды», в которой происходит размножение частиц в момент времени n. «Влияние среды» на процессы размножения выражается зависимостью закона распределения вероятностей $p_{ij}(x_n)$, $j=0, 1, \ldots$, числа непосредственных потомков каждой из частиц n-го поколения от x_n. При этом все частицы размножаются независимо.

Легко видеть, что

$$\mathbf{M} s^{Z_n^{\mathbf{x}}} = \mathcal{P}_{x_0}(\mathcal{P}_{x_1}(\ldots \mathcal{P}_{x_{n-1}}(s)\ldots))^i,$$

если $Z_0^{\mathbf{x}} = i$. Выпишем *многомерную* п. ф. сл. в. $Z_0^{\mathbf{x}}, \ldots, Z_n^{\mathbf{x}}$

$$\mathcal{P}^{(n)}(s_0, s_1, \ldots, s_n) = \sum_{\{i_0, i_1, \ldots, i_n\}} p_{Z_0^{\mathbf{x}}, \ldots, Z_n^{\mathbf{x}}}(i_0, \ldots, i_n)\, s_0^{i_0} \ldots s_n^{i_n} =$$

$$= \mathbf{M} s_0^{Z_0^{\mathbf{x}}} s_1^{Z_1^{\mathbf{x}}} \cdot \ldots \cdot s_n^{Z_n^{\mathbf{x}}} = \mathbf{M}\mathbf{M}(s_0^{Z_0^{\mathbf{x}}} \ldots s_n^{Z_n^{\mathbf{x}}} \,|\, Z_0^{\mathbf{x}}, \ldots, Z_{n-1}^{\mathbf{x}}) =$$

$$= \mathbf{M}(s_0^{Z_0^{\mathbf{x}}} \ldots s_{n-1}^{Z_{n-1}^{\mathbf{x}}} \mathbf{M}(s_n^{Z_n^{\mathbf{x}}} \,|\, Z_0^{\mathbf{x}}, \ldots, Z_{n-1}^{\mathbf{x}})) = \mathbf{M}(s_0^{Z_0^{\mathbf{x}}} \ldots s_{n-1}^{Z_{n-1}^{\mathbf{x}}} \mathcal{P}_{x_{n-1}}(s_n)^{Z_{n-1}^{\mathbf{x}}}) =$$

$$= \mathcal{P}^{n-1}(s_0, \ldots, s_{n-2}, s_{n-1}\, \mathcal{P}_{x_{n-1}}(s_n));$$

$$\mathcal{F}^{(n)}(s_0, \ldots, s_n) = (s_0 \mathcal{P}_{x_0}(s_1 \mathcal{P}_{x_1}(s_2 \mathcal{P}_{x_2}(\ldots s_{n-1} \mathcal{P}_{x_{n-1}}(s_n) \ldots))))^i,$$

если $Z_0^{\mathbf{x}} = i$.

Представим себе теперь, что задан некоторый случайный процесс $\xi = (\xi_0, \xi_1, \xi_2, \ldots)$ со значениями из $\mathscr{X}$, и рассмотрим ξ_n как *случайное состояние среды* размножения в момент времени n. Определим ветвящийся процесс Z_n^{ξ} в случайно изменяющейся среде ξ, построив по реализации $\xi = \mathbf{x}$ среды неоднородный ветвящийся процесс $Z_n^{\mathbf{x}}$. Далее предполагаем, что ξ есть последовательность независимых одинаково распределенных сл. в. Тогда $Z_n^{\xi} \equiv Z_n$ можно определить как марковскую цепь с переходными вероятностями

$$p_{ij} = \mathbf{M} p_{ij}(\xi_n), \quad i, j = 0, 1, 2, \ldots,$$

где предполагаем, что функции $p_{ij}(x)$ — измеримые, так что $p_{ij}(\xi_n)$ есть случайная величина. В самом деле, для многомерной п. ф. сл. в. $Z_0, Z_1, \ldots, Z_n$ имеем, как и выше,

$$\mathbf{M} s_0^{Z_0} \ldots s_n^{Z_n} = \mathbf{M}\mathbf{M}(s_0^{Z_0} \ldots s_n^{Z_n} | \xi_i, Z_i, i \leqslant n-1) =$$

$$= \mathbf{M}(s_0^{Z_0} \ldots s_{n-1}^{Z_{n-1}} \mathbf{M}(s_n^{Z_n} | \xi_i, Z_i, i \leqslant n-1))] =$$

$$= \mathbf{M}(s_0^{Z_0} \ldots s_{n-1}^{Z_{n-1}} \mathcal{P}_{\xi_{n-1}}(s_n)^{Z_{n-1}}) = \ldots =$$

$$= \mathbf{M}(s_0 \mathcal{P}_{\xi_0}(s_1 \mathcal{P}_{\xi_1}(\ldots s_{n-1} \mathcal{P}_{\xi_{n-1}}(s_n) \ldots)))^{Z_0}.$$

П. ф. под знаком математического ожидания можно записать в виде

$$\sum_{\{i_0, i_1, \ldots, i_n\}} p_{Z_0}(i_0) p_{i_0 i_1}(\xi_0) \ldots p_{i_{n-1} i_n}(\xi_n) s_0^{i_0} s_1^{i_1} \ldots s_{n-1}^{i_{n-1}}.$$

Применяя операцию математического ожидания (см. 10.15) и используя независимость сл. в. $Z_0, \xi_0, \xi_1, \ldots, \xi_n$, получаем

$$\mathbf{M} s_0^{Z_0} s_1^{Z_1} \ldots s_n^{Z_n} = \sum_{\{i_0, i_1, \ldots, i_{n-1}\}} p_{Z_0}(i_0) p_{i_0 i_1} \ldots p_{i_{n-1} i_n} s_0^{i_0} s_1^{i_1} \ldots s_{n-1}^{i_{n-1}}.$$

Учитывая, что манипуляции с условными математическими ожиданиями обоснованы полностью в § 9 лишь для дискретных сл. в., определим формально *ветвящийся процесс в среде из независимых одинаково распределенных величин* как марковскую цепь с переходными вероятностями вида

$$p_{ij} = \int_{-\infty}^{\infty} p_{ij}(x)\, dF(x), \quad i, j = 0, 1, 2, \ldots,$$

где $F(x)$ — некоторая вероятностная функция распределения.

Вводя последовательность $\xi_0, \xi_1, \xi_2, \ldots$ независимых сл. в. с ф. р. $F(x)$, имеем

$$p_{i_0i_1}p_{i_1i_2}\ldots p_{i_{n-1}i_n} = \mathbf{M}p_{i_0i_1}(\xi_0)\,\mathbf{M}p_{i_1i_2}(\xi_1)\ldots \mathbf{M}p_{i_{n-1}i_n}(\xi_{n-1}) =$$
$$= \mathbf{M}p_{i_0i_1}(\xi_0)\,p_{i_1i_2}(\xi_1)\ldots p_{i_{n-1}i_n}(\xi_{n-1}),$$

откуда получаем при $Z_0 = i_0$ многомерную п. ф. для $Z_1, \ldots, Z_n$ в виде

$$\mathbf{M}(s_0\mathcal{P}_{\xi_0}(s_1\mathcal{P}_{\xi_1}(\ldots s_{n-1}\mathcal{P}_{\xi_{n-1}}(s_n)\ldots)))^{i_0}$$

без обращения к свойствам условного математического ожидания.

Поставим вопрос о вероятности вырождения ветвящегося процесса в независимой одинаково распределенной среде. Положим

$$T(\omega) = n \Leftrightarrow Z_{n-1}(\omega) > 0, \quad Z_n(\omega) = 0$$

и $T(\omega) = \infty$ для тех ω, для которых $Z_n(\omega) > 0$ при всех n. Пусть для определенности $Z_0 = 1$. Тогда

$$\mathbf{P}(\omega : T(\omega) > n) = 1 - \mathcal{P}_{Z_n}(0) = \mathbf{M}(1 - \mathcal{P}_{\xi_0}(\mathcal{P}_{\xi_1}(\ldots \mathcal{P}_{\xi_{n-1}}(0)\ldots))).$$

Запишем очевидное неравенство

$$1 - \mathcal{P}_{\xi_0}(\mathcal{P}_{\xi_1}(\ldots \mathcal{P}_{\xi_{n-1}}(0)\ldots)) \leqslant \mathcal{P}'_{\xi_0}(1-)\,\mathcal{P}'_{\xi_1}(1-)\ldots \mathcal{P}'_{\xi_m}(1-), \quad m \leqslant n-1.$$

Введем сл. в. $X_i = \ln \mathcal{P}'_{\xi_i}(1-)$, $i = 0, 1, 2, \ldots$, и перепишем правую часть неравенства в виде

$$\exp(S_m), \quad S_m = X_0 + \ldots + X_m.$$

Для получения наиболее тесно прилегающей оценки сверху выберем $m = m_n(\omega)$ из условия минимума правой части неравенства, т. е.

$$S_{m_n(\omega)} = \min(S_0, S_1, \ldots, S_{n-1}) \equiv \mu_n.$$

Таким образом,

$$\mathbf{P}(\omega : T(\omega) > n) \leqslant \mathbf{M}\exp(\mu_n),$$

где $\mu_n = \mu_n(\omega)$ есть минимум случайного блуждания S_i, $i = 0, 1, 2, \ldots$, на отрезке $0 \leqslant i \leqslant n-1$. Отсюда видно, что если

$$\mathbf{P}(\omega : \varliminf_{n\to\infty} S_n(\omega) = -\infty) = 1,$$

то ветвящийся процесс — вырождающийся: $\mathbf{P}(\omega : T(\omega) < \infty) = 1$.

Рассмотрим более подробно частный случай процесса, предположив, что в зависимости от состояния среды размножение частиц идет по одному из двух геометрических распределений вероятностей: $p_{1j}(x) = \theta^j(1-\theta)$ либо $p_{1j}(x) = (1-\theta)^j\theta$, $j = 0, 1, 2, \ldots$. Так как $\mathcal{P}'_x(1)$ есть $\theta(1-\theta)^{-1}$ и $(1-\theta)\theta^{-1}$ соответственно, то сл. в.

$X_i = \ln \mathscr{P}'_{\xi_i}(1-)$ принимает значения $\alpha = \ln(\theta(1-\theta)^{-1}) = \ln\theta - \ln(1-\theta)$ и $-\alpha$. Пусть $\alpha > 0$ и

$$\mathbf{P}(\omega : X_i(\omega) = \alpha) = p, \quad 0 < p < 1.$$

В рассматриваемом случае блуждание $S_i = X_0 + \ldots + X_i$ лишь масштабом отличается от простого случайного блуждания с вероятностью p шага направо. При $p \leqslant 1/2$ нижний предел последовательности $S_n(\omega)$, $n = 0, 1, 2, \ldots$, равен $-\infty$ с вероятностью 1, так что процесс является вырождающимся, если

$$\mathbf{M}X_1 = \mathbf{M}\ln \mathscr{P}'_{\xi_1}(1-) = \alpha p - \alpha(1-p) = \alpha(2p-1) \leqslant 0.$$

Отметим, что в такой форме утверждение сохраняется и для процессов общего вида.

Воспользуемся тем, что (см. 14.12)

$$1 - \mathscr{P}(s) = (((1-s)\mathscr{P}'(1))^{-1} + \mathscr{P}''(1)/(2\mathscr{P}'(1)^2))^{-1}$$

для произвольной дробно-линейной п. ф. $\mathscr{P}(s)$ и что

$$\mathscr{P}''(1)/(2\mathscr{P}'(1)^2) = 1$$

в случае геометрического распределения. Отсюда получаем

$$1 - \mathscr{P}_{\xi_0}(\mathscr{P}_{\xi_1}(\ldots \mathscr{P}_{\xi_{n-1}}(s)\ldots)) = \left(\left((1-s)\prod_{i=0}^{n-1}\mathscr{P}'_{\xi_i}(1)\right)^{-1} + 1\right)^{-1},$$

$$\mathbf{P}(\omega : T(\omega) > n) = \mathbf{M}\left(\left(\prod_{i=0}^{n-1}\mathscr{P}'_{\xi_i}(1)\right)^{-1} + 1\right)^{-1} = \mathbf{M}\left(1 + \sum_{i=0}^{n-1} e^{-S_i}\right)^{-1}.$$

Если $\mathbf{M}X_1 = \alpha(2p-1) > 0$, т. е. $p > 1/2$, случайное блуждание $S_n \to \infty$ с вероятностью 1, при этом $S_n/n \to p$, так что с вероятностью 1

$$1 + \sum_{i=1}^{\infty} e^{-S_i} < \infty.$$

По теореме о монотонной сходимости 10.15 получаем

$$\lim_{n\to\infty} \mathbf{P}(\omega : T(\omega) > n) = \mathbf{M}\left(1 + \sum_{i=0}^{\infty} e^{-S_i}\right)^{-1} > 0.$$

Итак, при $\mathbf{M}X_1 > 0$ ветвящийся процесс в случайной среде с геометрическим распределением числа потомков имеет положительную вероятность невырождения. Этот результат сохраняется и в общем случае. Отметим еще, что условная вероятность невырождения до момента n при условии фиксации среды

$$\mathbf{P}(\{\omega : T(\omega) > n\} \mid \xi) = \left(1 + \sum_{i=0}^{n-1} e^{-S_i}\right)^{-1}$$

сходится при $n \to \infty$ к случайной величине

$$\zeta(\omega) = \left(1 + \sum_{i=0}^{\infty} e^{-S_i}\right)^{-1}$$

с вероятностью 1. Это означает, что почти все реализации случайной среды таковы, что условный процесс размножения (при условии фиксации среды), который представляет собой неоднородный ветвящийся процесс, имеет положительную (условную) вероятность невырождения.

14.39. Пример (дискретные процессы восстановления). В различных примерах марковских цепей мы встретились со следующим явлением: $p_{ij}^{(n)} \to p_j$ при $n \to \infty$ для любого состояния i, где p_j, $j \in \mathcal{J}$, есть некоторое распределение вероятностей. При этом мы говорили о сходимости марковской цепи к стационарному режиму. Каково бы ни было начальное состояние (или начальное распределение) такой марковской цепи, по прошествии длительного времени процесс становится близким к стационарному. Если распределение p_j, $j \in \mathcal{J}$, принять в качестве начального, то марковская цепь оказывается стационарным процессом: конечномерные распределения

$$\mathbf{P}(\omega : X_{n_k+n}(\omega) = j_k,\ k = 1, \ldots, m)$$

в этом случае не зависят от сдвига времени n. Стационарность процесса означает неизменность вероятностных закономерностей функционирования процесса на любом временном интервале. Не следует путать это свойство с однородностью марковской цепи — неизменностью условных вероятностных распределений

$$\mathbf{P}(\{\omega : X_{n_k+n}(\omega) = j_k,\ k = 1, \ldots, m\} \mid \{\omega : X_{n_0+n}(\omega) = j_0\}).$$

Как мы видели, необходимым условием сходимости марковской цепи к стационарному режиму является *возвратность*: с вероятностью 1 марковская цепь, выходящая из состояния j, в него возвращается, и притом бесконечно часто (для тех j, что $p_j > 0$). Мы видели, что если вероятность возвращения $f_{jj} < 1$, то возвращение в состояние j происходит конечное число раз с вероятностью 1, а потому $p_{jj}^{(n)} \to 0$ при $n \to \infty$.

Пусть марковская цепь такова, что для некоторого состояния j имеем $f_{jj} = 1$. Предположим, что это состояние j взято в качестве начального, и введем последовательные моменты возвращения в состояние j:

$$\tau_0 = 0 < \tau_1(\omega) < \tau_2(\omega) < \ldots .$$

Марковская цепь разбивается на независимые циклы

$$X_{\tau_i}, X_{\tau_i+1}, \ldots, X_{\tau_{i+1}}, \quad i = 0, 1, 2, \ldots,$$

между последовательными моментами возвращения в состояние j, эти циклы — точные вероятностные копии друг друга. Случайные моменты времени τ_i, $i=1, 2, \ldots$, являются *моментами восстановления* марковской системы в том смысле, что, начиная с любого τ_i, система функционирует независимо от прошлого и в точности по тем же вероятностным законам, как если бы τ_i было начальным моментом. Короче, в моменты τ_i система полностью восстанавливается до своего первоначального вида (регенерируется). Отвлекаясь от закономерностей функционирования процесса на *периоде регенерации* (τ_i, τ_{i+1}), свойственных марковской цепи, отметим, что процессы, обладающие указанным свойством восстановления, имеют важное прикладное значение и называются *регенерирующими*. Сходимость марковских цепей к стационарному режиму оказывается частным проявлением аналогичного свойства более широкого класса регенерирующих процессов.

Для любого $n>0$ положим

$$\nu_n(\omega)=k \Leftrightarrow \tau_k(\omega)\leqslant n, \quad \tau_{k+1}(\omega)>n.$$

Интервал времени от τ_{ν_n} до n имеет принципиальное значение при изучении предельного поведения распределения сл. в. X_n, $n\to\infty$, так как τ_{ν_n} есть последний перед n момент восстановления, и, следовательно, состояние марковской цепи в момент n полностью определяется ее поведением на интервале (τ_{ν_n}, n). Легко записать распределение вероятностей длины этого интервала:

$$\mathbf{P}(\omega : n-\tau_{\nu_n}(\omega)=t)=p_{jj}^{(n-t)} \sum_{s=t+1}^{\infty} f_{jj}^{(s)}, \quad t\leqslant n,$$

где $f_{jj}^{(s)}$ есть вероятность того, что марковская цепь, выходящая из состояния j, на s-м шаге возвратится в j впервые. Существование предела

$$\lim_{n\to\infty} p_{jj}^{(n)}=p_j>0$$

очевидно эквивалентно наличию предельного распределения последовательности сл. в. $n-\tau_{\nu_n}(\omega)$, $n=1, 2, \ldots$. При этом

$$\lim_{n\to\infty} \mathbf{P}(\omega : n-\tau_{\nu_n}(\omega)=t)=p_j \sum_{s=t+1}^{\infty} f_{jj}^{(s)}, \quad t=0, 1, 2, \ldots .$$

Так как

$$\sum_{t=0}^{\infty} \sum_{s=t+1}^{\infty} f_{jj}^{(s)}=\sum_{t=0}^{\infty} \mathbf{P}(\omega : \tau_1(\omega)>t)=\mathbf{M}\tau_1,$$

то в предположении, что предельное распределение сл. в. $n-\tau_{\nu_n}$, $n=1, 2, \ldots$, собственное, получаем $\mathbf{M}\tau_1<\infty$, $p_j=(\mathbf{M}\tau_1)^{-1}$.

Если цепь в начальный момент находится в некотором состоянии i, то, сохраняя обозначение ν_n для числа попаданий в состояние j до момента времени n, получаем

$$\mathbf{P}(\omega : n-\tau_{\nu_n}(\omega)=t)=p_{ij}^{(n-t)}\sum_{s=t+1}^{\infty} f_{jj}^{(s)}.$$

В предположении, что марковская цепь сходится к стационарному режиму

$(p_{ij}^{(n)}\to p_j,\ n\to\infty)$, имеем для сл. в. $n-\tau_{\nu_n}$, $n=1, 2, \ldots$,

то же предельное распределение, что и при $i=j$. Это утверждение сохраняется и для произвольного начального распределения. Точно так же из соотношения

$$\mathbf{P}(\omega : \tau_{\nu_n+1}(\omega)-n=t)=p_{ij}^{(n+t)}-\sum_{s=1}^{t-1} p_{ij}^{(n+s)}f_{jj}^{(t-s)}$$

(оно получается, если из всех путей, заканчивающихся в момент $n+t$ в состоянии j, отбросить пути, которые на временном участке от $n+1$ до $n+t-1$ попадают в состояние j) находим при $t=1, 2,\ldots$

$$\mathbf{P}(\omega : \tau_{\nu_n+1}(\omega)-n=t)\to p_j\sum_{s=t}^{\infty} f_{jj}^{(s)}.$$

Если же начальное распределение марковской цепи стационарное, то при любом $n=0, 1, 2, \ldots$

$$\mathbf{P}(\omega : \tau_{\nu_n+1}(\omega)-n=t)=\sum_{(i)} p_i\left(p_{ij}^{(n+t)}-\sum_{s=1}^{t-1} p_{ij}^{(n+s)}f_{jj}^{(t-s)}\right)=$$

$$=p_j-\sum_{s=1}^{t-1}\left(\sum_{(i)} p_i p_{ij}^{(n+s)}\right)f_{jj}^{(t-s)}=p_j\sum_{s=t}^{\infty} f_{jj}^{(s)}.$$

Итак, определив последовательные моменты

$$\tau_0(\omega)<\tau_1(\omega)<\tau_2(\omega)<\ldots$$

попадания марковской цепи в состояние j $(f_{jj}=1)$, получаем, что сл. в. $\Delta_i=\tau_i-\tau_{i-1}$, $i=0, 1, 2, \ldots$ $(\tau_{-1}=0)$, независимы, а при $i\geqslant 1$ одинаково распределены. В случае, когда марковская цепь сходится к стационарному режиму, сл. в. $n-\tau_{\nu_n}$ и $\tau_{\nu_n+1}-n$, $n=1, 2, \ldots$, имеют собственное предельное распределение. Для стационарной цепи сл. в. $\tau_{\nu_n+1}-n$, $n=0, 1, 2, \ldots$, одинаково распределены. Поскольку сл. в. ν_n+1 — момент остановки для последовательности сл. в. Δ_i, $i=0, 1, 2, \ldots$, то легко показать, что в случае стационарной цепи процесс $\tau_{\nu_n+k}-n$, $k=1, 2, \ldots$, является при любом n точной вероятностной копией исходного процесса. Описанные свойства последовательности сл. в. τ_n, $n=0, 1, 2, \ldots$, как мы увидим, имеют место для произвольного регенерирующего процесса с моментами восстановления τ_n, $n=0, 1, 2, \ldots$, и конечным средним периода регенерации.

Пусть Δ_i, $i=0, 1, 2, \ldots$, — произвольная последовательность положительных (кроме Δ_0, от которой требуем, чтобы $\Delta_0 \geqslant 0$) дискретных независимых сл. в., одинаково распределенных при $i \geqslant 1$. Последовательность сл. в. $\tau_n = \Delta_0 + \Delta_1 + \ldots + \Delta_n$, $n=0, 1, 2, \ldots$, называется *процессом восстановления*. Процесс восстановления при $\Delta_0 = 0$ называется *чистым*, а при $\mathbf{P}(\omega : \Delta_0(\omega) > 0) > 0$ — процессом с *запаздыванием*. Определим случайные моменты $\nu_n = \nu_n(\omega)$ условием $\nu_n = k \Leftrightarrow \tau_k \leqslant n$, $\tau_{k+1} > n$, k, $n=0, 1, 2, \ldots$. Принципиальную роль при изучении процессов восстановления играет *функция восстановления*

$$H(n) = \mathbf{M}\nu_n, \quad n=1, 2, \ldots, \quad H(0) = \mathbf{P}(\omega : \tau_0(\omega) = 0).$$

Из представления

$$\nu_n(\omega) = \sum_{k=0}^{\infty} I_{\{\omega : \tau_k(\omega) \leqslant n\}}(\omega)$$

получаем

$$H(n) = \sum_{k=0}^{\infty} \mathbf{P}(\omega : \tau_k(\omega) \leqslant n).$$

Заметим, что ввиду положительности случайных величин τ_i, $i \geqslant 1$, сумма справа содержит конечное число членов, и, кроме того, события $\{\omega : \tau_k(\omega) = n\}$, $k=0, 1, 2, \ldots$, не пересекаются, так что

$$\mathbf{P}(\omega : \tau_k(\omega) = n \text{ при некотором } k) =$$

$$= \sum_{k=0}^{\infty} \mathbf{P}(\omega : \tau_k(\omega) = n) = H(n) - H(n-1) \equiv h_n.$$

В случае, когда процесс восстановления представляет собой моменты возвращения марковской цепи в некоторое состояние j, h_n представляет собой вероятность попадания в момент n в состояние j. Если марковская цепь сходится к стационарному режиму, то

$$h_n \to (\mathbf{M}\Delta_1)^{-1}, \quad n \to \infty.$$

С некоторыми оговорками это утверждение сохраняется для произвольных процессов восстановления и носит название *теоремы восстановления*. Поясним связь этого утверждения с законом больших чисел для независимых одинаково распределенных величин. Из соотношения $\nu_n > m \Leftrightarrow \tau_m < n$ имеем в предположении $\mu = \mathbf{M}\Delta_1 < \infty$

$$\frac{\nu_n}{n} - \mu^{-1} > x \Leftrightarrow \nu_n > n(x + \mu^{-1}) \Leftrightarrow \frac{\tau_{[n(x+\mu^{-1})]}}{[n(x+\mu^{-1})]} < \frac{n}{[n(x+\mu^{-1})]}.$$

Но $\tau_n/n \xrightarrow{p} \mu$, так что при $x > -\mu^{-1}$, $x \neq 0$ и $n \to \infty$

$$\mathbf{P}(\omega : \nu_n(\omega)/n - \mu^{-1} > x) \to E\left((x+\mu^{-1})^{-1} - \mu\right),$$

где $E(x)$ есть ф. р. вероятностного распределения, сосредоточенного в нуле. Таким образом,

$$\nu_n/n \xrightarrow{p} \mu^{-1}, \quad n \to \infty.$$

Записав это соотношение в виде $\nu_n = n\mu^{-1} + o_p(1)n$, где множитель $o_p(1)$ мал по вероятности, получаем, что число восстановлений приблизительно равно длине интервала, деленной на среднюю длину периода восстановления. В предыдущем соотношении можно перейти к математическому ожиданию и получить *теорему восстановления в слабой форме* (см. 14.44)

$$H(n)/n = \mathbf{M}\nu_n/n \to \mu^{-1}, \ n \to \infty.$$

14.40. Задача. Предположим, что функция восстановления $H(n)$ удовлетворяет соотношению

$$h_n \equiv H(n) - H(n-1) \to \mu^{-1}, \ n \to \infty.$$

Вывести отсюда, что при $n \to \infty$

$$\mathbf{P}(\omega : \tau_{\nu_n+1}(\omega) - n = t) \to \mathbf{P}(\omega : \Delta_1(\omega) \geqslant t)\mu^{-1}, \quad t > 0,$$

$$\mathbf{P}(\omega : n - \tau_{\nu_n}(\omega) = t) \to \mathbf{P}(\omega : \Delta_1(\omega) \geqslant t+1)\mu^{-1}, \quad t \geqslant 0.$$

[$\mathbf{P}(\omega : \tau_{\nu_n+1}(\omega) - n = t) = \mathbf{P}(\omega : \tau_0(\omega) - n = t) +$

$$+ \sum_{k=0}^{\infty} \mathbf{P}(\omega : \tau_k(\omega) \leqslant n, \ \tau_{k+1}(\omega) - n = t).$$

Второе слагаемое в правой части преобразуем к виду

$$\sum_{k=0}^{\infty}\sum_{j=0}^{n} \mathbf{P}(\omega : \tau_k(\omega) = n - j, \ \Delta_{k+1}(\omega) = j + t) =$$

$$= \sum_{j=0}^{n}\sum_{k=0}^{\infty} \mathbf{P}(\omega : \tau_k(\omega) = n - j)\,\mathbf{P}(\omega : \Delta_{k+1}(\omega) = j + t) =$$

$$= \sum_{j=0}^{n-1} h_{n-j}\mathbf{P}(\omega : \Delta_1(\omega) = j + t) + \mathbf{P}(\omega : \tau_0(\omega) = 0)\,\mathbf{P}(\omega : \Delta_1(\omega) = n + t).$$

Переходя здесь к пределу при $n \to \infty$ и учитывая, что

$$\sum_{j=m}^{n} \mathbf{P}(\omega : \Delta_1(\omega) = j + t) \leqslant \mathbf{P}(\omega : \Delta_1(\omega) \geqslant m) < \varepsilon$$

для любого $\varepsilon > 0$ и достаточно больших m, а также, что $h_n \leqslant 1$, получаем из $h_n \to \mu^{-1}$

$$\mathbf{P}(\omega : \tau_{\nu_n+1}(\omega) - n = t) \to \mathbf{P}(\omega : \Delta_1(\omega) \geqslant t)\mu^{-1}, \quad n \to \infty.]$$

14.41. Задача. Предположим, что в процессе восстановления с конечным средним выполнено соотношение

$$\mathbf{P}(\omega : \Delta_0(\omega) = t) = \mathbf{P}(\omega : \Delta_1(\omega) \geqslant t)/\mathbf{M}\Delta_1 \quad t = 1, 2, \ldots.$$

Вычислить п. ф. последовательности $h_n = H(n) - H(n-1)$, $n = 1, 2, \ldots$, и убедиться, что $h_n = \text{const}$, $\mathbf{P}(\omega : \tau_{\nu_n+1}(\omega) - n = t) = \mathbf{P}(\omega : \Delta_0(\omega) = t)$.
[Положив $\mathscr{P}(s) = \mathbf{M}s^{\Delta_1}$, $\mu = \mathbf{M}\Delta_1$, имеем

$$Q(s) = \mathbf{M}s^{\Delta_0} = \sum_{j=0}^{\infty} p_{\Delta_0}(j) s^j = \mu^{-1} \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} p_{\Delta_1}(i) s^j =$$

$$= \mu^{-1} \sum_{i=0}^{\infty} p_{\Delta_1}(i) \sum_{j=0}^{i} s^j = s\mu^{-1}(1 - \mathscr{P}(s))/(1-s),$$

$$\sum_{n=1}^{\infty} h_n s^n = \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} p_{\tau_k}(n) s^n = \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} p_{\tau_k}(n) s^n =$$

$$= \sum_{k=0}^{\infty} \mathbf{M}s^{\tau_k} = \sum_{k=0}^{\infty} Q(s) \mathscr{P}(s)^k = Q(s)/(1 - \mathscr{P}(s))$$

при $0 < s < 1$. Подставляя в последнее выражение формулу для $Q(s)$ из предыдущего соотношения, получаем

$$\sum_{n=1}^{\infty} h_n s^n = \mu^{-1} s/(1-s),$$

откуда находим $h_n = \mu^{-1}$, $n = 1, 2, \ldots$. Далее (ср. 14.40)

$$\mathbf{P}(\omega : \tau_{\nu_{n+1}}(\omega) - n = t) = \mathbf{P}(\omega : \Delta_0(\omega) - n = t) +$$

$$+ \sum_{j=0}^{n-1} h_{n-j} p_{\Delta_1}(j+t) = \mu^{-1}\mathbf{P}(\omega : \Delta_1(\omega) \geqslant n+t) +$$

$$+ \mu^{-1}\mathbf{P}(\omega : t \leqslant \Delta_1(\omega) \leqslant n+t-1) = \mu^{-1}\mathbf{P}(\omega : \Delta_1(\omega) \geqslant t).]$$

14.42. Задача. Допустим, что состояния i, j марковской цепи $X_n(\omega)$, $n = 0, 1, \ldots$, таковы, что $f_{ij} = f_{jj} = 1$ (f_{kl} — вероятность когда-то попасть в l, выходя из k). Обозначим $\tau_m(\omega)$ последовательные моменты попадания марковской цепи в состояние j, $m = 0, 1, 2, \ldots$ (при $i = j$ получаем $\tau_0 = 0$). Предположим, что функция восстановления

$$H(n) = \sum_{k=0}^{\infty} \mathbf{P}(\omega : \tau_k(\omega) \leqslant n)$$

удовлетворяет соотношению

$$H(n)-H(n-1)\to(\mathbf{M}\Delta_1)^{-1},\ n\to\infty,\ \mathbf{M}\Delta_1=\mathbf{M}(\tau_1-\tau_0)<\infty.$$

Показать, что для любого состояния k марковской цепи, которое достижимо из j ($f_{jk}>0$), выполняется предельное соотношение

$$p_{ik}^{(n)}\to p_k^*,\quad n\to\infty,$$

где вероятности $p_k{}^*$ образуют стационарное распределение марковской цепи.

[Отметим, что из $f_{jk}>0$ в условиях задачи следует $f_{jk}=1$ (см. конец примера 14.35), а следовательно, и $f_{ik}=1$. Из соотношений

$$p_{ik}^{(n)}=\sum_{m=0}^{n} f_{ij}^{(m)} p_{jk}^{(n-m)},\quad \sum_{m=0}^{\infty} f_{ij}^{(m)}=f_{ij}=1$$

(где $f_{ij}^{(m)}$ есть вероятность попадания впервые в состояние j на m-м шаге, выходя из i) следует, что достаточно установить предельное соотношение

$$p_{jk}^{(n)}\to p_k^*,\quad n\to\infty.$$

Обозначим $q_{jk}^{(m)}$, $k\neq j$, вероятность попадания на m-м шаге в состояние k, выходя из j, не побывав до этого в состоянии j. Предполагая, что $X_0(\omega)=j$, очевидно, имеем

$$p_{jk}^{(n)}=q_{jk}^{(n)}+\sum_{l=1}^{\infty}\mathbf{P}(\omega:\tau_l(\omega)\leqslant n,\ \tau_{l+1}(\omega)>n,\ X_n(\omega)=k)=$$

$$=q_{jk}^{(n)}+\sum_{l=1}^{\infty}\sum_{m=0}^{n-1}\mathbf{P}(\omega:\tau_l(\omega)=n-m)\,q_{jk}^{(m)}=$$

$$=q_{jk}^{(n)}+\sum_{m=0}^{n-1}(H(n-m)-H(n-m-1))\,q_{jk}^{(m)}.$$

Заметим, что

$$\sum_{\{k:k\neq j\}} q_{jk}^{(m)}=\mathbf{P}(\omega:\Delta_1(\omega)>m),$$

$$q_{jk}\equiv\sum_{m=1}^{\infty} q_{jk}^{(m)}\leqslant\sum_{m=1}^{\infty}\mathbf{P}(\omega:\Delta_1(\omega)>m)=M\Delta_1-1<\infty,$$

$$\sum_{\{k:k\neq j\}} q_{jk}=\sum_{m=1}^{\infty}\sum_{\{k:k\neq j\}} q_{jk}^{(m)}=M\Delta_1-1.$$

Переходя в предыдущем соотношении к пределу при $n\to\infty$, получаем

$$\lim_{n\to\infty} p_{jk}^{(n)}=q_{jk}/\mathrm{M}\Delta_1, \quad k\neq j.$$

При $k=j$ имеем

$$p_{jj}^{(n)}=H(n)-H(n-1)\to 1/\mathrm{M}\Delta_1, \quad n\to\infty,$$

так что, полагая $q_{jj}=1$, имеем общую формулу для предельного распределения

$$p_k{}^*=q_{jk}/\mu_j,$$

где μ_j есть среднее время возвращения в состояние j.

Отметим, что из доказанного вытекает (см. 14.39), что для любого k, такого, что $p_k{}^*>0$ (т. е. $f_{jk}>0$) верно предельное соотношение

$$p_{jk}^{(n)}\to\mu_k^{-1},$$

откуда имеем равенство $\mu_k^{-1}=q_{jk}\mu_j^{-1}$.]

14.43*. Задача. Пусть Δ_i, $i=1, 2, \ldots$, — неотрицательные независимые одинаково распределенные случайные величины, $\mathrm{M}\Delta_1<<\infty$, $\tau_n=\Delta_1+\ldots+\Delta_n$, $n=1, 2, \ldots$, $\tau_0=0$, $\nu_t=\nu_t(\omega)$ — наибольшее значение индекса n, при котором $\tau_n\leqslant t$. Показать, что

$$\mathrm{M}\tau_{\nu_t+1}=\mathrm{M}\Delta_1\mathrm{M}(\nu_t+1).$$

[Введем сл. в.

$$X_i=\begin{cases}1, & \text{если } \tau_{i-1}\leqslant t,\\ 0, & \text{если } \tau_{i-1}>t.\end{cases}$$

Очевидно,

$$\tau_{\nu_t+1}=\Delta_1+\ldots+\Delta_{\nu_t+1}=\sum_{i=1}^{\infty}\Delta_i X_i,$$

$$\nu_t+1=X_1+\ldots+X_{\nu_t+1}=\sum_{i=1}^{\infty}X_i.$$

Применяя операцию математического ожидания и пользуясь возможностью переставить со знаком суммирования (см. 10.15), получаем

$$\mathrm{M}(\nu_t+1)=\sum_{i=1}^{\infty}\mathrm{M}X_i,\quad \mathrm{M}\tau_{\nu_t+1}=\sum_{i=1}^{\infty}\mathrm{M}(\Delta_i X_i)=\sum_{i=1}^{\infty}\mathrm{M}\Delta_i\mathrm{M}X_i=$$

$$=\mathrm{M}\Delta_1\sum_{i=1}^{\infty}\mathrm{M}X_i=\mathrm{M}\Delta_1\mathrm{M}(\nu_t+1),$$

где мы воспользовались независимостью сл. в Δ_i и X_i.]

14.44*. Задача. Пусть $H(t)$ — функция восстановления чистого процесса восстановления, $\mu<\infty$ — среднее время периода восстановления. Вывести из 14.43, что

$$H(t)/t\to\mu^{-1},\quad t\to\infty,$$

переходя к ограниченным сл. в. ($\Delta_i^{(c)}(\omega)=\Delta_i(\omega)$, если $\Delta_i(\omega)\leqslant c$ и $\Delta_i^{(c)}(\omega)=c$, если $\Delta_i(\omega)>c$).

[Обозначим $\nu_t^{(c)}$, $H^{(c)}(t)=M\nu_t^{(c)}$ соответствующие величины для процесса восстановления $\tau_n^{(c)}=\Delta_1^{(c)}+\ldots+\Delta_n^{(c)}$, $n=1, 2, \ldots$. Так как $\Delta_i^{(c)}\leqslant\Delta_i$, то $\nu_t^{(c)}\geqslant\nu_t$, и следовательно (см. 14.43),

$$H(t)/t=\mathbf{M}\nu_t/t\leqslant\mathbf{M}\nu_t^{(c)}/t=H^{(c)}(t)/t=\mathbf{M}\nu_t^{(c)}/t<\mathbf{M}(\nu_t^{(c)}+1)/t=$$

$$=\mathbf{M}\tau^{(c)}_{\nu_t^{(c)}+1}/(t\mathbf{M}\Delta_1^{(c)})\leqslant(t+c)/(t\mathbf{M}\Delta_1^{(c)}),$$

откуда

$$\overline{\lim_{t\to\infty}}\,H(t)/t\leqslant 1/\mathbf{M}\Delta_1^{(c)}$$

при любом $c>0$. Переходя к пределу при $t\to\infty$, находим

$$\overline{\lim_{t\to\infty}}\,H(t)/t\leqslant\mu^{-1}.$$

С другой стороны,

$$H(t)=\mathbf{M}(\nu_t+1)-1=\mu^{-1}\mathbf{M}\tau_{\nu_t+1}-1>\mu^{-1}t-1,$$

что и заканчивает доказательство.]

14.45*. Задача. Рассмотрим два независимых процесса восстановления τ_n, τ_n', $n=0, 1, 2, \ldots$, таких, что $\Delta_n\equiv\tau_n=\tau_{n-1}\overset{\mathscr{D}}{=}\tau_n'-\tau_{n-1}'\equiv\Delta_n'$, $n\geqslant1$ (сл. в. совпадают по распределению), распределение сл. в. τ_0 произвольно, а распределение сл. в. τ_0' таково, что процесс восстановления τ_n', $n=0, 1, 2, \ldots$, стационарен:

$$\mathbf{P}(\omega:\tau_0'(\omega)=t)=\mathbf{P}(\omega:\Delta_1(\omega)\geqslant t)/\mathbf{M}\Delta_1,\quad \mathbf{M}\Delta_1<\infty.$$

Введем случайный момент времени $N=N(\omega)$, полагая $N=n$ при $\tau_n=\tau_n'$, $\tau_k\neq\tau_k'$, $k<n$, и $N=\infty$, если таких n нет, и предположим, что $\mathbf{P}(\omega:N(\omega)<\infty)=1$. Образуем последовательность сл. в. (рис. 37)

$$\tau_0,\ \tau_1,\ \ldots,\ \tau_N,\ \tau_{N+1}',\ \tau_{N+2}',\ \ldots.$$

Показать, что полученная последовательность совпадает по распределению с процессом восстановления τ_n, $n=0, 1, 2, \ldots$. Вывести отсюда, что

$$(H(t+s)-H(t))-(H'(t+s)-H'(t))\to0,\quad t\to\infty,$$

для любого $s>0$; здесь $H(t)$ — функция восстановления процесса τ_n, $n=0, 1, 2, \ldots$, $H'(t)=t/\mathbf{M}\Delta_1$ — функция восстановления процесса τ_n', $n=0, 1, 2, \ldots$.

[Сл. в. $N(\omega)$ является, очевидно, моментом остановки последовательности (Δ_i, Δ_i'), $i=0, 1, 2, \ldots$, независимых пар сл. в., так что (ср. 11.3, 11.4) последовательность $(\Delta_{N+i}, \Delta'_{N+i})$, $i=1, 2, \ldots$, распределена так же, как последовательность (Δ_i, Δ_i'), $i=1, 2, \ldots$, и не зависит от начального отрезка (Δ_i, Δ_i'), $i\leqslant N$.

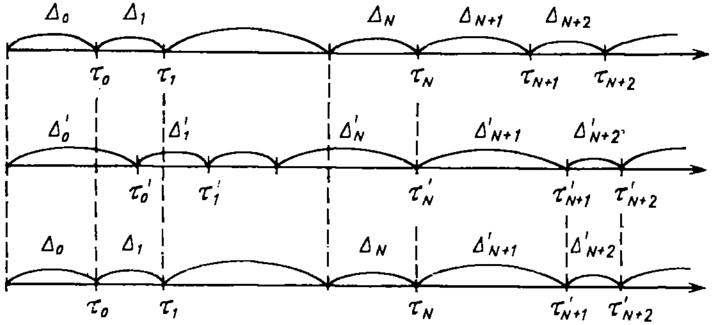

Рис. 37

Поэтому сл. в. $\Delta_0, \Delta_1, \ldots, \Delta_N, \Delta'_{N+1}, \Delta'_{N+2}, \ldots$ независимы и все, кроме Δ_0, распределены как Δ_1. Формальная выкладка состоит в вычислении распределения

$$\mathbf{P}(\omega : \tilde{\Delta}_i(\omega)=s_i, \quad i=1, \ldots, k),$$

где $\tilde{\Delta}_i(\omega)=\Delta_i(\omega)$ при ω, таких, что $N(\omega)\geqslant i$, и $\tilde{\Delta}_i(\omega)=\Delta_i'(\omega)$ при $N(\omega)<i$ и совершенно элементарна. Итак, последовательность

$$\tilde{\tau}_n=\tilde{\Delta}_0+\tilde{\Delta}_1+\ldots+\tilde{\Delta}_n, \quad n=0, 1, 2, \ldots,$$

совпадает по распределению с процессом восстановления τ_n, $n=0, 1, 2, \ldots$. Обозначим ν_t, ν_t' и $\tilde{\nu}_t$ соответственно наибольшее значение индекса n, при котором $\tau_n\leqslant t$, $\tau_n'\leqslant t$ и $\tilde{\tau}_n\leqslant t$, так что соответствующие функции восстановления записываются в виде

$$H(t)=\mathbf{M}\nu_t, \quad H'(t)=\mathbf{M}\nu_t', \quad \tilde{H}(t)=\mathbf{M}\tilde{\nu}_t,$$

причем $\tilde{H}(t)=H(t)$, $t\geqslant 0$. На множестве $\{\omega : \tau_N(\omega)\leqslant t\}$ имеем (рис. 37) $\tilde{\nu}_{t+s}(\omega)=\nu'_{t+s}(\omega)$, $s\geqslant 0$. Поскольку сл. в. N конечна с вероятностью 1, то также конечна сл. в. τ_N и

$$\mathbf{P}(\omega : \tau_N(\omega)\leqslant t)\to 1, \quad t\to\infty,$$

так что для п. в. ω $\tilde{\nu}_t(\omega)-\nu_t'(\omega)\to 0$ при $t\to\infty$. Следовательно, с вероятностью 1

$$(\tilde{\nu}_{t+s}(\omega)-\tilde{\nu}_t(\omega))-(\nu'_{t+s}(\omega)-\nu'_t(\omega))\to 0, \quad t\to\infty.$$

Взяв математическое ожидание от левой части

$$\mathbf{M}((\tilde{\nu}_{t+s}-\tilde{\nu}_t)-(\nu'_{t+s}-\nu'_t)),$$

применим теорему Лебега о мажорированной сходимости 10.24 (с учетом того, что из $\Delta_i \geqslant 1$, $\Delta_i' \geqslant 1$ имеем $\nu_{t+s}-\nu_t \leqslant s$, $\nu'_{t+s}-\nu_t' \leqslant s$) и получим требуемый результат.

Приведем еще одно доказательство основного результата:

$$H(t+s)-H(t)=\mathbf{M}(\nu_{t+s}-\nu_t)=\mathbf{M}((\nu_{t+s}-\nu_t)I_{\{\omega:\nu_t<N\}})+$$

$$+\mathbf{M}((\nu_{t+s}-\nu_t)I_{\{\omega:\nu_t\geqslant N\}})\equiv M_1+M_2,$$

$$M_1 \leqslant s\mathbf{M}I_{\{\omega:\nu_t<N\}}=s\mathbf{P}(\omega:\tau_N(\omega)>t)\to 0, \quad t\to\infty,$$

$$M_2=\mathbf{M}((\nu'_{t+s}-\nu'_t)I_{\{\omega:\nu_t\geqslant N\}})=\mathbf{M}(\nu'_{t+s}-\nu'_t)-$$

$$-\mathbf{M}((\nu'_{t+s}-\nu'_t)I_{\{\omega:\nu_t<N\}})=H'(t+s)-H'(t)-M_1',$$

$$M_1' \equiv \mathbf{M}((\nu'_{t+s}-\nu'_t)I_{\{\omega:\nu_t<N\}})\leqslant s\mathbf{P}(\omega:\tau_N(\omega)>t).]$$

14.46. Задача. Пусть сл. в. X_i, $i=1, 2, \ldots$, независимы, одинаково распределены, принимают целые значения, распределение вероятностей X_i симметрично ($\mathbf{P}(\omega:X_1(\omega)=k)=\mathbf{P}(\omega:X_1(\omega)=-k)$) и $\mathbf{M}X_1$ конечно ($\mathbf{M}X_1=0$). Показать, используя формулу обращения для характеристических функций 13.55, что

$$\sum_{n=0}^{\infty}\mathbf{P}(\omega:S_{2n}(\omega)=0)=\infty, \quad S_n=X_1+\ldots+X_n,$$

и, следовательно, марковская цепь $S_0=0$, S_1, S_2, ... с вероятностью 1 возвращается в начальное состояние.

[Ввиду симметрии распределения вероятностей сл. в. X_1 ее характеристическая функция

$$\psi(t)=\mathbf{M}e^{itX_1}=\mathbf{M}e^{-itX_1}=\overline{\psi}(t)=\psi(-t)$$

действительна и четна. По формуле обращения 13.55 запишем

$$\mathbf{P}(\omega:S_n(\omega)=0)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\psi(t)^n\,dt=\frac{1}{\pi}\int_0^{\pi}\psi(t)^n\,dt.$$

Применяя теорему о монотонной сходимости 10.15, получаем

$$\sum_{n=0}^{\infty}\mathbf{P}(\omega:S_{2n}(\omega)=0)=\frac{1}{\pi}\sum_{n=0}^{\infty}\int_0^{\pi}\psi(t)^{2n}\,dt=\frac{1}{\pi}\int_0^{\pi}\frac{dt}{1-\psi(t)^2}.$$

Так как $\mathbf{M}X_1=i\psi'(0)=0$, то

$$0\leqslant 1-\psi(t)=o(t), \quad t\to 0,$$

откуда вытекает, что интеграл справа обращается в ∞. Таким образом, для переходных вероятностей $p_{kl}^{(n)}$, k, $l=0, \pm 1, \pm 2, \ldots$, марковской цепи $S_j = X_1 + \ldots + X_j$, $j=0, 1, 2, \ldots$, $S_0=0$, имеем

$$\sum_{n=1}^{\infty} p_{00}^{(n)} \geqslant \sum_{n=1}^{\infty} p_{00}^{(2n)} = \infty,$$

откуда вытекает, что с вероятностью 1 происходит возвращение в начальное состояние (см. 14.35).]

14.47*. Задача. В условии 14.46 введем множество

$$K^+ = \{k > 0 : \mathbf{P}(\omega : X_1(\omega) = k) > 0\}$$

всех значений сл. в. X_1, принимаемых с положительной вероятностью. Обозначим L множество целых чисел, представимых в виде

$$m_1 k_1 + \ldots + m_r k_r, \quad k_i \in K^+, \quad i=1, \ldots, r,$$

при некоторых натуральных r, m_1, $\ldots$, m_r. Предположим, что наибольший общий делитель (НОД) чисел из K^+ равен 1. Показать, что множество K^+ содержит все натуральные числа, начиная с некоторого. Вывести отсюда, что марковская цепь $S_0=0$, S_1, S_2, $\ldots$ с вероятностью 1 попадает в любую целую точку.

[Пусть d есть НОД чисел из некоторого множества K^+. Расположив элементы из K^+ в возрастающую последовательность $0 < a_1 < a_2 < \ldots$, имеем, что d — делитель чисел a_1, a_2, $\ldots$, a_n при любом n. Тогда при некотором l d — НОД чисел $a_1, \ldots, a_l$. Действительно, обозначим d_n НОД чисел a_1, $\ldots$, a_n. Поскольку d_{n+1} делит d_n, $n=1, 2, \ldots$, то последовательность d_n, $n=1, 2, \ldots$, не возрастает и имеет предел, а следовательно, начиная с некоторого n, все члены последовательности d_n совпадают со своим пределом, т. е. с d.

Из алгоритма Евклида нахождения НОД вытекает, что НОД конечного набора чисел a_1, $\ldots$, a_l можно представить в виде

$$d = c_1 a_1 + \ldots + c_l a_l,$$

где c_1, $\ldots$, c_l — некоторые целые (возможно отрицательные) числа. В самом деле, при $l=2$, производя по алгоритму Евклида, деления с остатком (в предположении, что a_2 не является делителем a_1), получаем

$$a_1 = a_2 q_1 + r_2, \ 0 < r_2 < a_2, \ a_2 = r_2 q_2 + r_3, \ 0 < r_3 < r_2, \ \ldots,$$

$$r_{n-2} = r_{n-1} q_{n-1} + r_n, \ 0 < r_n < r_{n-1}, \ r_{n-1} = r_n q_n.$$

Легко видеть, что множество общих делителей пар чисел (a_1, a_2) (a_2, r_2), $\ldots$, (r_{n-2}, r_{n-1}) одно и то же, а следовательно, последний ненулевой остаток r_n является НОД чисел a_1, a_2. Подставим выражение для r_2 из первого равенства во второе, из полученного равенства выразим r_3 и подставим в третье и т. д. В результате получим искомое представление $d = c_1 a_1 + c_2 a_2$. В общем случае достаточно заметить, что НОД d_l чисел $a_1, \ldots, a_l$ совпадает с НОД чисел d_{l-1}, a_l.

Пусть теперь $d=1$, так что при некоторых целых $c_1, \ldots, c_l$ имеем

$$1=c_1a_1+\ldots+c_la_l.$$

Любое натуральное число m разделим с остатком на $s=a_1+\ldots+a_l$ и запишем

$$m=sq+r=(a_1+\ldots+a_l)q+r(c_1a_1+\ldots+c_la_l)=$$
$$=a_1(q+rc_1)+\ldots+a_l(q+rc_l),\quad 0\leqslant r<s.$$

Если m настолько велико, что $q+rc_i>0$, $i=1, \ldots, l$, то получаем, что $m\in L$.

Из сказанного вытекает, что при $n=|c_1|+\ldots+|c_l|$

$$\mathbf{P}(\omega : S_n(\omega)=1)=\mathbf{P}(\omega : S_n(\omega)=-1)>0,$$

а следовательно, марковская цепь S_n, $n=0, 1, 2, \ldots$, попадает в любую целую точку с положительной вероятностью. Так как вероятность возвращения в нулевое состояние $f_{00}=1$, то из $f_{0k}>0$ следует, что $f_{0k}=1$ (см. конец примера 14.35).]

14.48*. Задача. В условии 14.45 показать, что предположение $\mathbf{P}(\omega : N(\omega)<\infty)=1$ выполняется, если множество чисел

$$\{k : \mathbf{P}(\omega : \Delta_1(\omega)=k)>0\}$$

имеет НОД 1.

[Равенство $\tau_n(\omega)=\tau_n'(\omega)$ означает, что

$$-(\Delta_1-\Delta_1')-(\Delta_2-\Delta_2')-\ldots-(\Delta_n-\Delta_n')=\tau_0-\tau_0',$$

где слева стоит сумма $S_n=X_1+\ldots+X_n$, $X_i=\Delta_i'-\Delta_i$ независимых одинаково распределенных симметричных сл. в., к которым применимо утверждение 14.47. Следовательно,

$$\mathbf{P}(\omega : S_n(\omega)=\tau_0(\omega)-\tau_0'(\omega) \text{ при некотором } n\geqslant 1)=$$
$$=\sum_{\{k\}}\mathbf{P}(\omega : S_n(\omega)=k \text{ при некотором } n,\ \tau_0(\omega)-\tau_0'(\omega)=k)=$$
$$=\sum_{\{k\}}\mathbf{P}(\omega : S_n(\omega)=k \text{ при некотором } n)\cdot\mathbf{P}(\omega : \tau_0(\omega)-\tau_0'(\omega)=k)=1.]$$

Задачи 14.45—14.48 приводят к следующей *теореме восстановления*. Пусть d есть НОД множества чисел

$$\{k : \mathbf{P}(\omega : \Delta_1(\omega)=k)>0\},$$

обозначим $l_d(I_{a,b})$ количество точек целочисленной решетки с шагом $d(0, \pm d, \pm 2d, \ldots)$, содержащихся в промежутке $I_{a,b}$. Тогда при $t\to\infty$ и любом $s>0$ имеет место соотношение

$$H(t+s)-H(t)-l_d((t, t+s])/\mathbf{M}\Delta_1\to 0,$$

где в случае $\mathbf{M}\Delta_1=\infty$ дробь считается равной нулю.

При $d=1$ сформулированное утверждение вытекает из 14.45, 14.48 и 14.41. При $d>1$ достаточно перейти к сл. в. $\tilde{\Delta}_i=\Delta_i/d$, $i=0, 1, 2, \ldots$.

Из теоремы восстановления выводится *теорема о предельном поведении переходных вероятностей* $p_{ij}^{(n)}$ марковской цепи. Именно, пусть $f_{ij}=f_{jj}=1$, $\mu_j\leqslant\infty$ есть среднее время возвращения в состояние j. Обозначим d_j период состояния j — НОД значений времени возвращения в j, имеющих положительную вероятность. Если $d_j=1$, то из 14.42 имеем

$$p_{ij}^{(n)}\to 1/\mu_j, \quad n\to\infty,$$

для любого i, такого, что $f_{ij}=1$, а сумма предельных вероятностей $p_j^*=1/\mu_j$ по всем таким j равна 1. Если период $d_j>1$, то аналогично имеем

$$p_{ij}^{(nd_j)}\to d_j/\mu_j, \quad n\to\infty.$$

Литература

1. Боровков А. А. Теория вероятностей. Изд. 2. М.: Наука, 1986.
2. Гнеденко Б. В. Курс теории вероятностей. Изд. 5. М.: Наука, 1974.
3. Климов Г. П. Теория вероятностей и математическая статистика. М. Изд-во МГУ, 1983.
4. Козлов М. В., Прохоров А. В. Введение в математическую статистику М.: Изд-во МГУ, 1987.
5. Колмогоров А. Н. Основные понятия теории вероятностей. Изд. 2. М. Наука, 1974.
6. Лоэв М. Теория вероятностей. М.: ИЛ, 1962.
7. Мешалкин Л. Д. Сборник задач по теории вероятностей. М.: Изд-во МГУ, 1963.
8. Невё Ж. Математические основы теории вероятностей. М.: Мир, 1969.
9. Партасарати К. Введение в теорию вероятностей и теорию меры. М.: Мир, 1983.
10. Петров В. В. Суммы независимых случайных величин. М.: Наука, 1972.
11. Прохоров А. В., Ушаков В. Г., Ушаков Н. Г. Задачи по теории вероятностей. Основные понятия. Предельные теоремы. Случайные процессы. М.: Наука, 1986.
12. Прохоров Ю. В., Розанов Ю. А. Теория вероятностей. Изд. 2. М. Наука, 1987.
13. Розанов Ю. А. Теория вероятностей, случайные процессы и математическая статистика. М.: Наука, 1985.
14. Севастьянов Б. А. Ветвящиеся процессы. М.: Наука, 1971.
15. Севастьянов Б. А. Курс теории вероятностей и математической статистики. М.: Наука, 1982.
16. Тутубалин В. Н. Теория вероятностей. М.: Изд-во МГУ, 1972.
17. Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1, 2. М.: Мир, 1984.
18. Хеннекен П. Л., Тортра А. Теория вероятностей и некоторые ее приложения. М.: Наука, 1974.
19. Чжун Кай-лай. Однородные цепи Маркова. М.: Мир, 1964.
20. Чистяков В. П. Курс теории вероятностей. Изд. 3. М.: Наука, 1987.
21. Ширяев А. Н. Вероятность. М.: Наука, 1980.

Список обозначений и сокращений

$n!=n\cdot(n-1)\cdot\ldots\cdot2\cdot1$, $0!=1$

$C_n^k=n!/(k!(n-k)!)$

$N_{n,k}=C_n^{(n+k)/2}$ при $|k|\leqslant n$, k одной четности с n, $N_{n,k}=0$ при остальных n и k

$p_n(k;\ p)=C_n^k p^k(1-p)^{n-k}$ — биномиальное распределение (см. § 1, 2)

$A\subseteq B$ — включение множества A в B

$A\cup B$ — объединение множеств

$A\cap B$, AB — пересечение множеств

$\overline{A}$ — дополнение множества $A\subseteq\Omega$ в Ω , $A\ominus B=A\overline{B}$

$\omega\in A$, $\omega\overline{\in}A$ — принадлежность, не принадлежность элемента множеству

$\varnothing$ — пустое множество

$\mathbf{P}(A)$, $\mathbf{P}(\omega:\ldots)$ — вероятность события A, вероятность события, описываемого после двоеточия

$p(\omega)$ — вероятность элементарного события ω

сл. в. — случайная величина

$X=X(\omega)$, $Y=Y(\omega)$, ... — случайные величины

$\mathbf{X}=(X_1,\ \ldots,\ X_n)$, $\mathbf{Y}=(Y_1,\ \ldots,\ Y_m)$ — случайные векторы или наборы сл. в.

$p_X(x)$, $p_{\mathbf{X}}(\mathbf{x})=p_{X_1,\ldots,X_n}(x_1,\ \ldots,\ x_n)$ — распределение вероятностей дискретной сл. в., совместное распределение вероятностей набора дискретных сл. в.

$\mathscr{X}_1\times\mathscr{X}_2\times\ldots\times\mathscr{X}_n$ — прямое произведение множеств $\mathscr{X}_i$, $i=1,\ \ldots,\ n$ (см. § 2) $\mathscr{X}^n=\mathscr{X}\times\mathscr{X}\times\ldots\times\mathscr{X}$ n раз, $\mathscr{X}^\infty=\mathscr{X}\times\mathscr{X}\times\ldots$

$I_{a,b}$ — общее обозначение промежутков числовой прямой с включенными либо невключенными концами $a\geqslant-\infty$, $b\leqslant\infty$

I_A — индикатор множества A

$\mathscr{A}$, $\mathscr{B}$, $\mathscr{D}$, $\mathscr{M}$, $\mathscr{U}$ — классы подмножеств из Ω (начиная с § 3)

$\mathscr{F}_X(x)$, $\mathscr{F}_{\mathbf{X}}(\mathbf{x})=\mathscr{F}_{X_1,\ldots,X_n}(x_1,\ \ldots,\ x_n)$ — функция распределения сл. в X, совместная функция распределения сл. в $X_1,\ \ldots,\ X_n$ (см. § 3)

ф. р. — функция распределения

$f_X(x)$, $f_{\mathbf{X}}(\mathbf{x})=f_{X_1,\ldots,X_n}(x_1,\ \ldots,\ x_n)$ — плотность распределения сл. в X, со- совместная плотность сл. в. $X_1,\ \ldots,\ X_n$ (см. § 3)

$X_{(1)}\leqslant X_{(2)}\leqslant\ldots\leqslant X_{(n)}$ — порядковые статистики сл. в. $X_1,\ \ldots,\ X_n$ (см. 3.4, 4.17), $X_{kn}\equiv X_{(k)}$

$p(k,\ \lambda)$ — распределение Пуассона (см. 3.11)

$\Gamma(\alpha)$ — гамма-функция Эйлера (см. § 3)

$\ln x$ — логарифм натуральный

$\exp(x)\equiv e^x$ — экспонента

$a_n\sim b_n$ — предел отношения a_n/b_n равен 1

$a_n\asymp b_n$ — $c<a_n/b_n<C$ для всех n и некоторых $0<c<C<\infty$

$a_n=o(1)$ — предел последовательности a_n равен 0

$a_n=O(1)$ — последовательность a_n ограничена

$a_n=o(b_n)$, $a_n=O(b_n)$ — $a_n/b_n=o(1)$, $a_n/b_n=O(1)$

$\varphi(x)$ — плотность стандартного нормального распределения (см. § 3)
$\det A$ — определитель матрицы A
$\mathcal{P}_X(t)$ — производящая функция сл. в. X (см. § 4)
п. ф. — производящая функция
R, R^n, R^∞ — действительная прямая, n-мерное арифметическое пространство, пространство последовательностей действительных чисел
$\prod_{i=1}^{k} x_i$ — произведение чисел $x_1, \ldots, x_k$
$\mathbf{P}(A|B)$, $\mathbf{P}_B(A)$ — условная вероятность события A при условии события B (см. § 5)
$p_{X|Y}(x|y)$, $p_{\mathbf{X}\,\mathbf{Y}}(\mathbf{x}|\mathbf{y})$ — условные распределения дискретных сл. в. (см. § 5)
p_{ij}, $p(j|i)$ — переходные вероятности марковской цепи (см. § 5)
$f_{X|Y}(x|y)$, $f_{\mathbf{X}\,\mathbf{Y}}(\mathbf{x}|\mathbf{y})$ — условные плотности распределений (см. § 5)
$\mathcal{B}$, $\mathcal{B}^n$ — класс борелевских множеств прямой, n-мерного арифметического пространства (см. § 6)
$\mathbf{M}$ — символ математического ожидания
$\mathbf{D}$ — символ дисперсии
X_+, X^+, X_-, X^-, f^+, f_- — положительные и отрицательные части функций X, f (см. § 7, 10)
$\mathbf{R}_\mathbf{X}$ — матрица ковариаций случайного вектора $\mathbf{X}$ (см. § 8);
$\mathbf{M}(X|A)$, $\mathbf{M}_A X$ — условное математическое ожидание сл. в. X при условии события A (см. § 9)
$\mathbf{M}(X|Y)$, $\mathbf{M}(X|\mathcal{A})$ — условное математическое ожидание сл. в. X при условии сл. в. Y, при условии алгебры событий $\mathcal{A}$ (см. § 9)
$\overline{\lim\limits_{n\to\infty}}\, a_n = \inf\limits_n \sup\limits_{k\geqslant n} a_k$ — верхний предел последовательности a_n
$\underline{\lim\limits_{n\to\infty}}\, a_n = \sup\limits_n \inf\limits_{k\geqslant n} a_k$ — нижний предел последовательности a_n
$\xrightarrow{\mathcal{D}}$, $\xrightarrow{p}$ — символы сходимости по распределению, по вероятности (см. § 13)
НОД — наибольший общий делитель
$g(a-) = \lim\limits_{x\to a-} g(x)$ — предел функции $g(x)$ при x, стремящемся к a слева
$g(+\infty)$, $g(-\infty)$ — предел функции $g(x)$ на бесконечности

Учебное издание

Козлов Михаил Васильевич

ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ В ПРИМЕРАХ И ЗАДАЧАХ

Зав. редакцией *С. И. Зеленский*
Редактор *Л. А. Николова*
Художественный редактор *Б. С. Вехтер*
Технический редактор *В. В. Макарова*
Корректоры *М. И. Эльмус, М. А. Мерецкова*

ИБ 3284

Сдано в набор 24.02.89. Подписано в печать 14.02.90.
Формат 60×90/16. Бумага тип. № 2.
Гарнитура литературная. Высокая печать.
Усл. печ. л. 21,5 Уч.-изд. л. 23,14 Тираж 5000 экз.
Заказ 28. Изд. № 582 Цена 1 р. 10 к.

Ордена «Знак Почета» издательство Московского университета.
103009, Москва, ул. Герцена, 5/7.
Типография ордена «Знак Почета» изд-ва МГУ.
119899, Москва, Ленинские горы