

С.Н. Дворяткина, Л.Н. Ляхов

**ЛЕКЦИИ
ПО КЛАССИЧЕСКОЙ
ТЕОРИИ ВЕРОЯТНОСТЕЙ**

Учебное пособие

**МОСКВА
URSS
2012**

Дворяткина Светлана Николаевна, Ляхов Лев Николаевич

Лекции по классической теории вероятностей. — М.: Книжный дом "ЛИБРОКОМ", 2012. — 181 с.

В основе данного курса лекций лежит классический подход, заключающийся в представлении материала с нарастающей глубиной заключений, что обеспечивает одновременно и высокий научный уровень, и простоту его восприятия. Вводимые понятия объяснены с точки зрения истории развития вопроса, что делает чтение лекций увлекательным занятием.

Курс лекций может быть использован так же в качестве учебного пособия преподавателями, студентами и аспирантами, изучающими физико-математические, информационные, технологические, инженерные и другие специальности.

Научный редактор: зав. кафедрой уравнений в частных производных и теории вероятностей Воронежского государственного университета, доктор физ.-мат. наук, проф. А.В. Глушко

Рецензенты: зав. кафедрой технической кибернетики и автоматического регулирования ВГУ, доктор тех. наук, проф. Г.И. Лозгачев;

доцент каф. математического анализа и элементарной математики ЕГУ им. И.А. Бунина, канд. пед. наук С.В. Щербатых

Содержание

Введение	6
1 Случайные события и их вероятности	13
§ 1 Пространство элементарных исходов	13
§ 2 Алгебра событий	16
2.1 Равновозможные исходы. Классическое определение вероятности	19
2.2 Применение элементов комбинаторики . . .	22
2.3 Геометрическая вероятность	26
2.4 Статистическая вероятность	29
§ 3 Аксиоматика теории вероятностей	32
§ 4 Условные вероятности. Независимость	38
4.1 Условная вероятность и теоремы умножения	38
4.2 Независимость событий	40
4.3 Формула полной вероятности. Формулы Байеса	43
§ 5 Последовательность испытаний	45
5.1 Понятие цепи Маркова	46
5.2 Схема испытаний Бернулли	49
5.3 Предельные теоремы в схеме испытаний Бернулли	52
2 Случайные величины. Функции распределения	61
§ 6 Случайные величины	64
6.1 Функция распределения вероятности	65

6.2	Плотность распределения вероятности . . .	70
§ 7	Числовые характеристики случайных величин . .	80
7.1	Математическое ожидание	80
7.2	Моменты. Дисперсия и среднее квадратиче- ское отклонение	84
7.3	Мода. Медиана	87
§ 8	Основные законы распределения	87
8.1	Равномерное распределение	87
8.2	Нормальное распределение	90
8.3	Показательное распределение	95
8.4	Биномиальное распределение	96
8.5	Распределение Пуассона	97
8.6	Геометрическое распределение	99
§ 9	Системы случайных величин	102
9.1	Понятие системы случайных величин . . .	102
9.2	Функция распределения системы $X[2]$. . .	104
9.3	Плотность вероятности системы $X[2]$. . .	106
9.4	Зависимые и независимые случайные вели- чины	109
9.5	Моменты, математическое ожидание, дис- персия системы $X[2]$	111
9.6	Ковариация случайных величин. Коэффи- циент корреляции	113
§ 10	Предельные теоремы теории вероятностей	115
10.1	Неравенство Чебышева	117
10.2	Теорема Чебышева	119
10.3	Закон больших чисел	121
10.4	Усиленный закон больших чисел	124
3	Элементы теории случайных процессов	128
§ 11	Введение в теорию случайных процессов	131
11.1	Определение случайного процесса	131
11.2	Классификация случайных процессов . . .	132
§ 12	Математический аппарат дискретных марковских цепей	136

12.1	Классификация марковских случайных процессов	137
12.2	Вероятностные характеристики цепей Маркова	138
12.3	Примеры цепей Маркова	143
12.4	Классификация состояний цепей Маркова .	146
12.5	Предельная теорема для конечных цепей Маркова	151
§ 13	Дискретный случайный марковский процесс . . .	152
13.1	Вероятностные характеристики марковских процессов	152
13.2	Система дифференциальных уравнений Колмогорова	154
13.3	Эргодические свойства однородных марковских случайных процессов	157
13.4	Пуассоновский процесс	160
13.5	Процесс чистого рождения	164
13.6	Процесс рождения и гибели	167
Дополнение №1		170
	δ -распределение Дирака	170
	Дифференцирование распределений	173
Дополнение №2. Таблицы		176
	Значения плотности нормального распределения .	176
	Значения функции Лапласа	177
	Значения функции вероятности распределения Пуассона	178
	Значения плотности и функции показательного распределения	179
Список литературы		180

Введение

Лишь случай — закон, а закон всегда случаен.

Эрик Мария Ремарк

При исследовании многих проблем, научных, социальных, производственных, военных и даже бытовых, часто приходится встречаться с особым типом явлений, которые принято называть случайными. Примерами таких случайностей могут служить траектории полета самолетов, случайные появления преград на железнодорожных путях, случайные отказы навигационных приборов и прочие. **Случайность** — это альтернатива закономерности. Закономерные (или детерминированные) явления изучаются в рамках классических научных теорий. Возникает вопрос, можно ли изучать случайности? Согласно логике и теории познания ответ очевидный — конечно, нет. Интуиция и угадывание исключаются. Ответ неочевидный — можно, но только в том случае, когда в случайностях присутствует закономерность.

Теория вероятностей — *математическая наука, изучающая закономерности в случайных явлениях*. Базовым понятием этой теории является *событие* — всякий исход некоего опыта (испытания), наблюдаемого исследователем. Классическими примерами являются события из области азартных игр (игры в монетку, кости, карты) — выпадение любой грани игральной кости, выпадение "орла" или "решки" и другие. Меры случайных событий, применяемые на бытовом уровне, хорошо известны (шанс, возможность, процент удачи, все равно и др.) Мерой события на математическом языке является *вероятность*. Это одна из числовых характеристик возможности появления случайного события, которая меняется от 0 до 1. Известны испытания, поставленные в одних и тех же условиях, которые всегда приводят к определенному событию (подброшенный камень всегда падает

на землю). Такие события называются достоверными и их вероятность равна 1. А известны опыты, которые показывают, что события никогда не наступят (монета, брошенная на стол, никогда не станет на ребро). Такие события называются невозможными, и их вероятность равна 0. Таким образом, случайные события всегда лежат между двумя крайностями — быть достоверными и невозможными, следовательно, их вероятностная мера лежит между 0 и 1.

Вопросы вероятностного характера возникали во многих сферах человеческой деятельности. Но главным стимулом развития теории вероятностей служили задачи, которые возникали в статистике, практике страховых обществ, при обработке результатов астрономических наблюдений. Большинство первых задач теории вероятностей было связано с азартными играми. Сегодня эти задачи не представляют никакой практической значимости, но с методической точки зрения — актуальны и полезны, так как позволяют избежать ошибок при построении вероятностных моделей и изучении вероятностных характеристик. В XVII веке появляются изящные решения отдельных задач, связанных с исследованием вероятности появления одного из *равновозможных*, исключающих друг друга событий. Самые простые примеры таких равновозможных событий дают игры в монетку, игральную кость, карты. Поэтому выделение в опыте (игре, процессе) набора простейших (элементарных) равновозможных исходов оказывалось наиболее важным, поскольку они сразу дают вероятность в виде величины обратной количеству всех равновозможных исходов. Но как оказалось, вопрос о равновозможности не обязательно простой. Наиболее поучительной и хорошо описанной в современной литературе является *задача де Мерре* [2], (обобщения и многие тонкости которой можно найти в книге Ю.Неймана [13]). История этой задачи такова. В конце семнадцатого века один французский вельможа шевалье де Мерре (известный своим современникам как любитель азартных игр, и в частности, игры в кости), заинтересовался математическими возможностями прогнозирования удачных ставок. Он заметил, что при бросании трех игральных костей комбинация, дающая в

сумме 11 очков, появляется чаще, чем 12, хотя вероятность появления обеих сумм казалась ему одинаковой, поскольку и ту, и другую комбинации можно получить одним и тем же числом способов:

$$\begin{aligned}
 11 &= 6 + 4 + 1 = 6 + 3 + 2 = 5 + 5 + 1 = \\
 &= 5 + 4 + 2 = 5 + 3 + 3 = 4 + 4 + 3; \\
 12 &= 6 + 5 + 1 = 6 + 4 + 2 = 6 + 3 + 3 = 5 + 5 + 2 = \\
 &= 5 + 4 + 3 = 4 + 4 + 4.
 \end{aligned}
 \tag{*}$$

Проведя длинную серию опытов и убедившись, что его математические рассуждения не согласуются с эмпирическими выводами, де Мере шлет гневное послание своему современнику, знаменитому философу, математику и механику Б. Паскалю (1623–1662), где обвинил науку и ученых в полной оторванности от мира и бессмысленности научного труда.

Блез Паскаль решил эту загадку следующим образом. Он предположил, что при бросании игральных костей комбинации очков (a,b,c) , (a,c,b) , (b,a,c) , (b,c,a) , (c,a,b) , (c,b,a) , где a , b , c — фиксированные числа, следует считать разными и равновероятными, т.е. равновероятными следует считать (взаимоисключающие) исходы, описываемые *упорядоченными* тройками чисел, а не исходы, описываемые просто тройками чисел, как это предполагалось в (*). Ясно, что если верно предположение Б. Паскаля, то количество комбинаций выпавших очков на трех игральных костях, определяемых тремя фиксированными числами, зависит от количества неодинаковых цифр в этой тройке. Например, комбинация из цифр 6,4,1 может выпасть 6 различными способами: $(6,4,1)$, $(6,1,4)$, $(1,4,6)$, $(1,6,4)$, $(4,1,6)$, $(4,6,1)$; из набора цифр, содержащих две одинаковые цифры — 3 способами, например: $(5,5,1)$, $(5,1,5)$, $(1,5,5)$; и наконец, набор, содержащий все одинаковые числа, например, $(4,4,4)$, может выпасть только одним способом. Теперь мы предоставим читателю возможность пересчитать в (*) количество различных "равновозможных" исходов для суммы 11 и для суммы 12. Результат, соответственно — 27 и 25, что и позволило Б. Паскалю объяснить парадокс де Мере.

Из этой истории мы должны сделать два очень важных вывода.

Первый: для одной задачи можно построить различные "вероятностные модели", и в каждой модели будут получены неодинаковые ответы на один и тот же вопрос. При этом совершенно необходимо ответить и на другой вопрос: какая из этих "моделей" справедлива (или справедливее)? Этим и обусловлено то, что современные курсы теории вероятностей начинаются с построения "вероятностного пространства" или некоторой вероятностной модели, в рамках которой справедливы соответствующие результаты исследования данной вероятностной задачи.

Второй: случайные события и случайные процессы подчинены неким объективным законам, которые удастся наблюдать только при достаточно большом количестве опытов. Эти законы такие же объективные, как закон всемирного тяготения или закон сохранения энергии, массы, вещества и т.д..

Второй вывод, по-видимому, сделали многие ученые, волей судьбы вынужденные заниматься задачами "о возможности". Примерами таких задач являлись задачи о подсчете количества исходов при бросании игральных костей, задачи о разделении ставки между двумя игроками, если игра прерывалась до выигрыша одним из игроков и др.. Так число различных исходов при бросании трех игральных костей было впервые определено в 965 году французским епископом В. Виболдом, но результатам опыта он придавал религиозную трактовку. Первое решение задачи о разделении ставки между двумя игроками предложил Лука Пачоли в 1494 году. Частные случаи правильного решения этой задачи были предложены позже Дж. Кардано (1501–1575). Следует упомянуть, что Дж. Кардано и Н. Тарталья были правильно решены многие задачи, связанные с бросанием игральных костей и выпадением того или иного числа очков. Однако верное решение задачи о разделении ставки игроков и описанного выше парадокса де Мере было впервые найдено в 1654 году Б. Паскалем (1623–1662) и П. Ферма (1601–1665) в ходе их знаменитой переписки. Наиболее же полный анализ задачи был проведен

Г. Галилеем (1564–1642) в работе "Рассуждения об игре в кости" (опубликованная только в 1718 г.). Дальнейшее формирование основ теории вероятностей, основных ее методов и теоретико-вероятностных понятий осуществлялось при исследовании частных задач и вопросов в практике статистики и страхования в работах Э. Галлея (1656–1742 гг.), Х. Гюйгенса (1629–1695), Д. Граунта (1620–1674), В. Петти (1623–1687).

Новый этап в истории теории вероятностей начался с исследований Я. Бернулли (1654–1705). основополагающие открытия в области теории вероятностей были изложены в работе "Искусство предположений" (1713 г.), где он доказывает первую предельную теорему (закон больших чисел), которая в дальнейшем послужит основой всех исследований о закономерностях появления случайных событий в массовых явлениях. Тем самым теория вероятностей получила важное практическое приложение и стала отдельной научной дисциплиной. Название новой науки впервые было предложено Паскалем, а в употребление вошло в 1718 году в связи с изданием книги Муавра "Учение о случаях". Непосредственным развитием закона больших чисел Я. Бернулли являются предельные теоремы, доказанные А. Муавром (1667–1754). К основным заслугам А. Муавра в области теории вероятностей можно также отнести вывод нормального закона распределения и разработку на основе азартных игр аналитического аппарата — теории возвратных последовательностей, которая была продолжена Л. Эйлером, Ж. Лагранжем и П. Лапласом.

"Классический" этап развития теории вероятностей завершается работами П.С. Лапласа (1749–1828). Были заново доказаны предельные теоремы, получены важные результаты в теории ошибок наблюдений, которые нашли широкие приложения в обработке результатов экспериментов. Но, справедливости ради, надо отметить, что в естествознании по-прежнему господствовали теории, порождаемые исследованием дифференциальных уравнений, а не вероятностные методы.

Классическая вероятность посвящена изучению множества событий, которые можно представить как объединение некото-

рого набора равновозможных и взаимоисключающих друг друга событий. Основные законы классической теории были открыты к середине XIX века, но оставался главный вопрос, имеет ли эта наука практическое применение. То, что *теория вероятностей*—математическая наука с огромными приложениями в физике, химии, генетике, биологии, социологии, медицине, экономике, стало ясно лишь в начале XX века, благодаря усилиям русских ученых П.Л.Чебышева (1821–1894), А.А. Маркова (1856–1922), А.Н. Колмогорова (1903–1987) и др.

П.Л. Чебышев ввел в рассмотрение случайную величину, доказал закон больших чисел в общей форме. А.А. Марков продолжает работу своего учителя П.Л. Чебышева. Он восполнил пробел в доказательстве основной предельной теоремы, тем самым дал полное и строгое доказательство этой теоремы в общих условиях, доказал предельные теоремы для последовательности зависимых величин, что привело к схеме "испытаний, связанных в цепь" (теория цепей Маркова). А.Н. Колмогоров построил систему аксиоматического обоснования теории вероятностей. Аксиоматический подход позволил распространить методы равновозможных на произвольные события и в ряде случаев объяснить, каким образом можно получить разные ответы одной задачи. Кроме того, А.Н. Колмогоровым введена аксиома непрерывности, позволившая рассматривать бесконечное (счетное) число исходов.

Надо отметить, что некоторые теоремы теории вероятностей весьма необычны по сравнению со всеми утверждениями других математических дисциплин. Несколько иронизируя, можно сказать, что теория вероятностей содержит утверждения, справедливые с некоторой вероятностью, в то время как другие математические дисциплины содержат утверждения, справедливые с вероятностью либо 1, либо 0. Выступая в московском математическом обществе, известный американский математик Дж. Дуб заметил в этой связи: "Всем специалистам по теории вероятностей хорошо известно, что математика представляет собой часть теории вероятностей" [11]. Далее читатель увидит, что с точки зрения "закона больших чисел" эта шутка выглядит всего лишь

"некоторым преувеличением".

Сегодня теория вероятностей продолжает бурно развиваться, в ней появляются новые направления исследований, которые представляют значительный общетеоретический и прикладной интерес.

Основные объекты изучения теории вероятностей:

- 1) случайное событие и его вероятность;
- 2) случайная величина и функции распределения;
- 3) случайный процесс и его вероятностная характеристика.

В данном пособии рассматриваются эти три части, разумеется, лишь в рамках программы для технических специальностей университетов и технологических вузов. При составлении учебного пособия использовались известные учебники "Курс теории вероятностей" В.П. Чистякова [19], "Курс теории вероятностей и математической статистики" Б.В. Севастьянова [16] и классический учебник по теории вероятностей для математических специальностей университетов и педагогических институтов "Курс теории вероятностей" Б.В. Гнеденко [4]. В настоящем пособии используются δ -функционалы Дирака для задания плотности случайных величин, заданных дискретно. В связи с тем, что теория функционалов не входит в программу многих технических вузов, в приложении 2 дается краткое описание δ -функционалов Дирака и производных от разрывных функций (типа функций Хевисайда) на языке функционалов.

В заключении хотелось бы отметить, что сегодня вопрос изучения случайностей предстает во всей ясности как завещание целой эпохи. Его решение является и исходной точкой, и перспективой — задачей для будущих поколений.

Глава 1

Случайные события и их вероятности

§ 1. Пространство элементарных исходов

Определение 1.1. Пусть результатом некоторого эксперимента может быть одно из событий A_1, A_2, \dots, A_n , причем в результате испытания одно из этих событий обязательно произойдет (полная группа событий) и эти события не могут произойти вместе (несовместные). Тогда набор событий $\{A_j\}_{j=1}^n$ называется пространством элементарных событий и обозначается $\Omega = \{A_j\}_{j=1}^n$.

Число n называется размерностью пространства элементарных исходов.

Примеры.

◇ 1) Единичное подбрасывание монеты. Возможные исходы испытания — выпадение "орла" или выпадение "решки". Тогда $\Omega = \{O; R\}$. Размерность $n = 2$.

◇ 2) Подбрасывание игральной кости один раз. Возможные исходы испытания — количество выпавших очков. $\Omega = \{1; 2; 3; 4; 5; 6\}$. Размерность $n = 6$.

◇ 3) Информационная система кодирует сигнал с помощью m -мерных векторов, координатами которых служат числа 0 и 1. Рассматривая такой вектор как случайное событие, построить в частных случаях $m = 1, m = 2, m = 3$ пространство элементарных событий и в общем случае определить его размерность.

Решение. Пусть $m = 1$. Тогда $\Omega = \{0; 1\}$. Размерность $n = 2$.

Пусть $m = 2$. Тогда $\Omega = \{(0, 0); (0, 1); (1, 0); (1, 1)\}$. Размерность $n = 4$.

Пусть $m = 3$. Тогда

$$\Omega = \{(0, 0, 0); (0, 0, 1); (0, 1, 0); (1, 0, 0); (1, 1, 0); (1, 0, 1); (0, 1, 1);$$

$(1, 1, 1)\}$. Размерность $n = 8$.

Покажем, что в общем случае пространство элементарных исходов состоит из 2^m векторов (событий). Применим метод математической индукции. Как мы уже убедились, наше предположение справедливо для $m = 1, 2, 3$. Предположим, что при $m = k$ пространство элементарных исходов состоит из 2^k векторов (событий). В каждый из таких векторов добавим одну координату (в начало) и заполним эту координату сначала нулем (при этом получим 2^k векторов), а затем единицей (получим еще 2^k векторов). Нетрудно видеть, что таким образом мы перебрали всевозможные комбинации $(k + 1)$ -мерных векторов из нулей и единиц. Их общее число $2^k + 2^k = 2^k 2 = 2^{k+1}$, что и доказывает предположение об объеме пространства элементарных исходов.

◇ 4) Работа сервера электронной почты. Возможные исходы — количество сообщений (корреспонденций), поступивших по электронной почте в течение определенного времени. Практически количество вызовов ограничено, но удобнее считать эту случайную величину неограниченной. Тогда простран-

ство элементарных исходов — бесконечная последовательность $\{0, 1, 2, 3, \dots\}$. Заметим, что в этом случае часто в качестве пространства элементарных событий можно рассматривать множество ступенчатых функций времени, каждая из которых описывает возможную реальную ситуацию следующим образом: она постоянна между моментами поступления корреспонденций и совершает единичный скачок в момент поступления одного вызова на *e-mail*. При этом мы наблюдаем и другие элементарные события — время поступления сообщения, которое принимает произвольное значение на соответствующем отрезке времени, и поэтому весь этот отрезок принимается за пространство элементарных событий $\Omega = \{[t_0, t_1]\}$.

◇ 5) Пусть информационная система кодирует стандартную информацию следующим образом. Некоторый квадрат на плоскости с декартовыми координатами разбит на части прямыми параллельными осям координат $x = x_j$, $y = y_i$, причем любой точке в данном прямоугольнике соответствует одна единственная фраза или ключевое слово. Исход испытания — точка на плоскости, где введена декартова система координат xOy . Пространством элементарных событий является бесконечное множество упорядоченных наборов (x, y) : $\Omega = \{(x, y), x \in [a, b], x \neq x_j, y \in [c, d], y \neq y_i\}$.

◇ 6) Работа годографа, фиксирующего движение частицы, совершающей броуновское движение в определенном промежутке времени. В этом случае пространство элементарных событий состоит из бесконечного (несчетного) множества негладких вектор-функций $\Omega = \{\bar{\xi} = (x(t), y(t), z(t))\}$.

◇ 7) Полет самолета. Возможные исходы — траектории полета самолета в заданном воздушном коридоре. Тогда пространство элементарных событий — это множество гладких линий в данном "коридоре", которые описываются набором дифференцируемых вектор-функций $(x(t), y(t), z(t))$.

§ 2. Алгебра событий

В реальном испытании, кроме взаимоисключающих исходов, составляющих пространство элементарных событий, можно указать другие события, каждое из которых оказывается некоторым подмножеством пространства элементарных событий. Это приводит к возможности использования теории множеств для введения алгебраических операций над событиями.

Пусть для некоторого испытания определено пространство элементарных событий Ω . Любое подмножество из Ω будем называть **событием**. Элементарные события, составляющие пространство элементарных событий Ω , будем обозначать ω или ω_i , а события (объединение ω) будем обозначать большими латинскими буквами A, B, C и т.д..

Событие A влечет за собой событие B ($A \subset B$), если при наступлении события A обязательно наступает и событие B .

Если событие A влечет за собой событие B , а событие B влечет за собой событие A , то события A и B называются эквивалентными или равными ($A=B$).

Суммой событий $A+B$ (объединение событий A и B) называется событие, состоящее в том, что произошло или событие A , или событие B (рис.1).

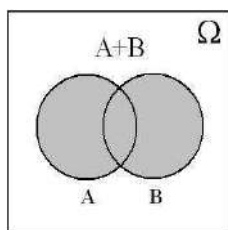


Рис.1

Произведением AB (пересечение событий A и B) будем называть событие, состоящее в том, что произошло и событие A , и событие B (рис.2).

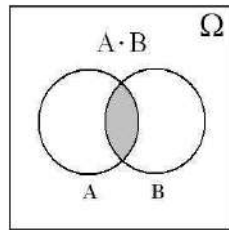


Рис.2

Разностью $A - B$ событий называется событие, заключающееся в том, что произошло событие A , но не произошло B (рис.3). В частности, если события A и B не могут произойти вместе (несовместны), то $A - B = A$.

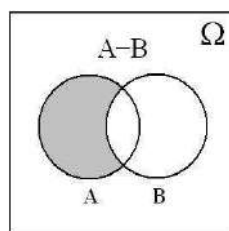


Рис.3

Пространство элементарных исходов Ω будем называть **до-стоверным** событием, а пустое множество \emptyset — **невозможным** событием.

Например, при бросании монеты вы приняли за пространство элементарных исходов $\Omega = \{ \text{„орел“}, \text{„решка“} \}$, тогда все события типа „монета не упала“, „встала на ребро“, „провалилась сквозь землю“ и др. считаются невозможными.

Событие $\bar{A} = \Omega - A$ называется **противоположным** событию A (читается "не A ") (рис. 4).

События A и B называются **несовместными**, если

$$AB = \emptyset.$$

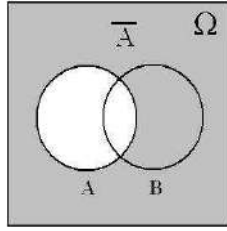


Рис.4

События A_1, A_2, \dots, A_n называются **попарно несовместными**, если для любых $i \neq j$, где $1 \leq i, j \leq n$, события A_i и A_j несовместны.

Пусть Ω — произвольное пространство элементарных событий, а \mathcal{U} — объединение некоторых подмножеств множества Ω . Класс подмножеств \mathcal{U} множества Ω называется **алгеброй событий**, если

- 1) $\Omega \in \mathcal{U}$;
- 2) $A \in \mathcal{U}, \Rightarrow \overline{A} \in \mathcal{U}$;
- 3) $A \in \mathcal{U}, B \in \mathcal{U} \Rightarrow \begin{cases} AB \in \mathcal{U}, \\ A + B \in \mathcal{U}, \\ A - B \in \mathcal{U}. \end{cases}$

Из определения алгебры \mathcal{U} вытекает, что $\emptyset \in \mathcal{U}$, поскольку $\emptyset = A - A, \forall A \in \mathcal{U}$.

Простейшими примерами алгебр служат *минимальная алгебра* $\mathcal{U} = \{\Omega, \emptyset\}$ и *максимальная алгебра*, которая включает в себя множество всевозможных подмножеств, которые могут быть составлены из элементов множества Ω , включающего в качестве одного из элементов само множество Ω .

♦ **Пример 1.** Пусть $\Omega = \{x : -\infty < x < \infty\}$ — числовая прямая. Рассмотрим систему подмножеств \mathcal{U}_0 на прямой из сплошных промежутков типа интервалов, полуинтервалов и отрезков. Это множество не является алгеброй, поскольку объединение множеств с "дыркой" на прямой (например, $[0; 1] \cup [2; 4]$) не принадлежит этой системе подмножеств. Дополним \mathcal{U}_0 всеми конечными суммами множеств из \mathcal{U}_0 . Полученная система мно-

жеств \mathcal{U} является алгеброй.

Часто приходится иметь дело с множеством Ω бесконечной размерности. В этом случае вводится понятие σ -алгебры (борелевой алгебры).

Определение 2.1. Алгебра событий \mathcal{U} называется σ -алгеброй, если из того, что каждый элемент из последовательности событий $\{A_j\}_{j=1}^{\infty}$ принадлежит \mathcal{U} , следует, что

$$\bigcap_{j=1}^{\infty} A_j \in \mathcal{U}, \quad \bigcup_{j=1}^{\infty} A_j \in \mathcal{U}.$$

♦ **Пример 2.** Множество всех рациональных чисел и множество всех действительных чисел — σ -алгебры.

Определение 2.2. Элементы σ -алгебры \mathcal{U} , заданной на множестве Ω , называются *случайными событиями*.

2.1. Равновозможные исходы. Классическое определение вероятности

Пусть пространство элементарных исходов $\Omega = \{\omega_i\}_{i=1}^n$ состоит из конечного числа n равновозможных исходов. И пусть \mathcal{U} это некоторая σ -алгебра.

Равновозможность элементарных событий позволит ввести функцию вероятности ¹ P на пространстве элементарных исходов по правилу $P(\omega_i) = \frac{1}{n}$ для произвольного $i = 1, 2, \dots, n$. Тогда произвольное событие $A \in \mathcal{U}$ представляет собой некоторое подмножество множества Ω :

$$A = \{A_{i_k}\}_{k=1}^m = \sum_{k=1}^m \omega_{i_k}, \quad m \leq n.$$

¹Термин "вероятность" впервые употребил Б. Паскаль в переписке с П. Ферма. В своем письме, датированном 1654 годом, он писал: "Замечу сразу же, что степень возможности (уверенности) события я назвал вероятностью. В своей теории я исхожу из того основного предположения, что каждому событию, наступление которого зависит от случая, можно поставить в соответствие определенное число, заключенное между 0 и 1, в качестве его вероятности. В дальнейшем я всюду буду пользоваться термином вероятность для обозначения числа, обозначающего степень уверенности".

Поэтому вероятность $P(A)$ события A равна отношению числа несовместных элементарных событий, составляющих событие A , к общему числу элементарных событий в Ω .

Определение 2.1.1. *Формула*

$$P = \frac{m}{n}, \quad (2.1.1)$$

где n — число возможных исходов, а m — число исходов, при которых событие A произойдет (благоприятных исходов), называется **классическим определением вероятности**².

Докажем справедливость следующих свойств вероятности, введенной по формуле (2.1.1.).

1. $P(A) \geq 0$.
2. $P(\Omega) = 1$.
3. Если события A и B несовместны, то $P(A + B) = P(A) + P(B)$.
4. $P(\bar{A}) = 1 - P(A)$.
5. $P(\emptyset) = 0$.
6. Если $A \subset B$, то $P(A) \leq P(B)$.
7. $0 \leq P(A) \leq 1$.

Д о к а з а т е л ь с т в о.

Свойство 1). Дробь m/n не может быть отрицательной.

Свойство 2). Так как событию Ω благоприятствуют все n возможных результатов испытаний, то $P(\Omega) = n/n = 1$.

Свойство 3). Пусть $\Omega = \{\omega_i\}_{i=1}^n$, $A = \{\omega_{i_k}\}_{k=1}^{s_1}$, $B = \{\omega_{j_k}\}_{k=1}^{s_2}$, $s_1 \neq s_2 \leq n$. Тогда

$$P(A + B) = \frac{s_1 + s_2}{n} = \frac{s_1}{n} + \frac{s_2}{n} = P(A) + P(B).$$

²Классическое определение вероятности вошло в науку в результате исследований многих ученых (Д. Кардано, Х. Гюйгенс, Я. Бернулли, П. Монмор и др.). В течение долгих лет это понятие корректировалось и совершенствовалось. Так согласно Я. Бернулли, вероятность есть "степень уверенности и относится к достоверности как часть к целому". Определение классической вероятности, используемое сегодня, точнее метод ее вычисления, дал П. Лаплас. Применяя современную аксиоматику А.Н. Колмогорова, можно легко вывести классическую формулу вероятности. Действительно, так как события ω_i — несовместные ($\omega_i \omega_j = \emptyset$, $i \neq j$), то по АЗ, получаем: $P(A) = P(\sum_{k=1}^m \omega_{i_k}) = \sum_{k=1}^m P(\omega_{i_k}) = \sum_{k=1}^m \frac{1}{n} = \frac{m}{n}$.

Свойство 4). Поскольку $A + \bar{A} = \Omega$, то по свойству 2: $P(A + \bar{A}) = 1$, так как A и \bar{A} несовместны, то по свойству 3 имеем: $P(A + \bar{A}) = P(A) + P(\bar{A})$.

Свойство 5). Действительно, $P(\Omega + \emptyset) = 1 = P(\Omega)$, а поскольку события Ω и \emptyset несовместны, то $P(\Omega) + P(\emptyset) = P(\Omega)$, $\Rightarrow P(\emptyset) = 0$.

Свойство 6). Если $A \subset B$, то $B = A + \bar{A} \cdot B$, тогда из свойств 3 и 1 получим:

$$P(B) = P(A + \bar{A} \cdot B) = P(A) + P(\bar{A} \cdot B) \geq P(A).$$

Свойство 7). В самом деле, для любого события A имеют место соотношения: $\emptyset \subset A + \emptyset = A = A \cdot \Omega \subset \Omega$. По свойству 6 имеем: $0 = P(\emptyset) \leq P(A) \leq P(\Omega) = 1$.

◇ **Пример 1 (теорема сложения).** Пусть $\Omega = \{\omega_i\}_{i=1}^n$, $A = \{\omega_{i_k}\}_{k=1}^{s_1}$, $B = \{\omega_{j_k}\}_{k=1}^{s_2}$, $AB = \{\omega_{q_k}\}_{k=1}^{s_3}$, $s_1, s_2, s_3 \leq n$.

Доказать, что

$$P(A + B) = P(A) + P(B) - P(AB).$$

Р е ш е н и е.

$$P(A + B) = \frac{s_1 + s_2 - s_3}{n} = \frac{s_1}{n} + \frac{s_2}{n} - \frac{s_3}{n} = P(A) + P(B) - P(AB).$$

Определение 2.1.2. Событие A , которое произошло при условии, что произошло событие B называется условным событием и обозначается A/B (читается "А при условии В"). Вероятность условного события называется условной вероятностью и обозначается $P(A/B)$.

◇ **Пример 2 (теорема умножения).** Пусть $\Omega = \{\omega_i\}_{i=1}^n$, $A = \{\omega_{i_k}\}_{k=1}^{s_1}$, $B = \{\omega_{j_k}\}_{k=1}^{s_2}$, $AB = \{\omega_{q_k}\}_{k=1}^{s_3}$, $s_1, s_2, s_3 \leq n$.

Доказать, что

$$P(AB) = P(A) P(B/A) = P(B/A) P(A).$$

Р е ш е н и е.

$$P(A/B) = \frac{s_3}{s_2} = \frac{s_3/n}{s_2/n} = \frac{P(AB)}{P(B)}, \implies P(AB) = P(A/B) P(B).$$

2.2. Применение элементов комбинаторики

При решении задач с равновозможным пространством элементарных событий необходимо создавать модели задач, позволяющие каким-либо образом пересчитать возможные и благоприятствующие исходы. При этом оказывается очень полезным знание формул комбинаторики, которые в ряде случаев позволят формализовать решение задач.

Применение формул комбинаторики рассмотрим на соответствующих примерах.

1. Количество комбинаций элементов из различных групп. Есть r групп элементов (произвольной природы: столы, стулья, лошади и т.д.), в каждой из которых n_i -элементов (т.е. $1 \leq i \leq r$). Из каждой группы выбирается один элемент, и составляется комбинация $\bar{a} = (a_1, a_2, \dots, a_r)$ так, что на i -ом месте стоит элемент из i -ой группы. Создается множество таких комбинаций. Равенство $\bar{a} = \bar{b}$ означает, что $a_i = b_i$. Сколько различных комбинаций в этом множестве?

Ответ: $N = n_1 n_2 \dots n_r$.

Ответ получается методом математической индукции: при $r = 2$ утверждение справедливо (расположить элементы на ортогональных прямых в целочисленных координатах и провести прямые, параллельные соответствующей оси, количество пересечений — $n_1 n_2$), предположим справедливость утверждения для $(r - 1)$ -ой групп, добавляем r -ую группу и рассматриваем только две группы, в первой из которых элементов $n_1 \dots n_{r-1}$, а во второй — n_r элементов, и применяем уже проверенную формулу для вычисления количества комбинаций для двух групп.

В частности, если в каждой из этих групп количество элементов одинаково, то $N = n^r$.

♦ **Пример 1.** Бросаются три игральные кости. Определить вероятность максимального количества суммарных очков.

Мы должны учесть парадокс де Мере, т.е. потребовать, чтобы пространство элементарных исходов Ω состояло из упорядоченных троек (a, b, c) , где каждое число может принять любое значение от 1 до 6. Ясно, что объем этого пространства элементарных исходов равен числу $N = 6^3 = 216$. Событие A — выпало 18 очков, т.е. комбинация $(6, 6, 6)$. Считая все элементы пространства Ω равновероятными, получим $p(A) = \frac{1}{216}$.

2. Выбор с возвращением и без. Пусть имеется группа из n элементов. Случайно вынимается один элемент, фиксируется, возвращается в группу, перемешивается, снова достается один элемент, фиксируется и т.д. r раз. Такая ситуация называется "выбор с возвращением". Получаем наборы типа $\bar{a} = (a_1, a_2, \dots, a_r)$. Два таких набора считаются разными, если хотя бы на одном месте стоят разные элементы. Сколько разных наборов можно составить таким выбором (с возвращением)?

Ответ: $N = n^r$.

Ответ получается из первой задачи, поскольку эту задачу можно интерпретировать так: имеется r групп, в каждой из которых n элементов.

Теперь рассмотрим "выбор без возвращения". Количество всевозможных исходов в этом случае называется "*числом размещений*".

Пусть имеется группа из n элементов. Случайно вынимается один элемент, фиксируется, не возвращается в группу, снова достается один элемент, фиксируется, не возвращается в группу и т.д. r раз. Затем процесс повторяется. В результате получаем наборы типа $\bar{a} = (a_1, a_2, \dots, a_r)$. Два таких набора считаются разными, если хотя бы на одном месте стоят различные элементы. Сколько разных наборов можно составить таким выбором (без возвращения)?

Ответ: $A_n^r = n(n-1)(n-2) \dots (n-r+1)$.

Получается из задачи 1 (количество комбинаций элементов из различных групп) следующей интерпретацией: в первой группе n элементов, во второй — $(n-1)$ элементов, в третьей $(n-2)$ элементов и т.д., в r -ой группе — $(n-(r-1))$ элементов.

Число A_n^r называется числом *размещений из n элементов*

по r .

В частности при $r = n$ получаем число $A_n^n = P_n = n!$, которое называется числом *перестановок из n элементов*.

♦ **Пример 2.** Пусть r студентов случайным образом рассаживаются по n вагонам электрички, $r \leq n$. Предположим, что каждый из студентов выбирает свой вагон случайно и попадает в любой из вагонов равновероятно. Какова вероятность, что все они попадут в разные вагоны?

Пространство элементарных исходов создадим из событий:

k -ый студент попал в i -ый вагон. Число таких событий

$N = n^r$ — выбор (вагона) без возвращения, поскольку студенты могут попасть и в один вагон. Однако число комбинаций, при которых удовлетворяется требование задачи, это выбор без возвращения: если первый студент попал в k -ый вагон, то второй выбирает среди $n - 1$ оставшихся вагонов, а третий выбирает среди $n - 2$ оставшихся и т.д.. По классической формуле искомая вероятность

$$p = \frac{n^r}{n(n-1) \dots (n-r+1)}.$$

3. Число сочетаний. Поставим вопрос: каким числом способов можно сочетать m элементов, выбираемых из данных различных n ($m \leq n$) элементов. Методом математической индукции доказывается формула

$$C_n^m = \frac{n!}{m!(n-m)!},$$

определяющая число всевозможных сочетаний из n элементов по m . Эти числа хорошо известны, т.к. представляют собой коэффициенты бинома Ньютона. Отметим следующее, часто необходимое при вычислениях свойство:

$$C_n^m = C_n^{n-m}.$$

Более общим является случай, когда совокупность из n элементов разбивается на фиксированное число k групп, причем фиксировано и число элементов в каждой группе, скажем, в

каждой группе с номером i содержится ровно n_i элементов. В остальном элементы группируются произвольно. В этом случае число различных сочетаний из n элементов по n_1, n_2, \dots, n_r , $n = n_1 + n_2 + \dots + n_r$ есть число

$$N = \frac{n!}{n_1! n_2! \dots n_k!},$$

называющееся числом *перестановок с повторяющимися элементами*.

♦ **Пример 3.** В ящике содержится 10 одинаковых деталей, помеченных номерами от 1 до 10. Наудачу извлечены 6 деталей. Найти вероятность того, что среди отобранных окажутся: а) деталь 1; б) деталь 1, деталь 2.

Р е ш е н и е.

а) Пространство элементарных событий Ω составим из всевозможных наборов из 6 деталей, т.е. 6 из 10. Объем этого пространства равен числу сочетаний из 10 по 6, т.е. C_{10}^6 . За алгебру в этом примере достаточно взять множество подмножеств из 6 элементов, добавив пустое и само множество из 10 деталей. События $\omega \in \Omega$ естественно считать равновероятными. Тогда остается посчитать количество элементарных исходов, при котором произойдет интересующее нас событие: среди отобранных 6 деталей есть деталь 1. Но тогда остальные пять деталей в этой группе имеют другие номера. Число таких исходов равно числу сочетаний из оставшихся 9 деталей (кроме детали 1) по 5, т.е. числу C_9^5 . Искомая вероятность определяется теперь по классическому определению вероятности

$$p = \frac{m}{n} = \frac{C_9^5}{C_{10}^6} = 0,6.$$

б) Пространство элементарных исходов то же, что и в предыдущей задаче: C_{10}^6 . Число благоприятных исходов (среди отобранных есть две фиксированные детали 1 и деталь 2, и, следовательно, оставшиеся 4 детали могут иметь любые другие номера) равно числу сочетаний из 8 по 4, т.е. числу C_8^4 . Искомая

вероятность

$$P = \frac{m}{n} = \frac{C_8^4}{C_{10}^6} = \frac{1}{3}.$$

♦ **Пример 4.** В партии из N деталей имеется n стандартных. Наудачу отобраны m деталей. Найти вероятность того, что среди отобранных ровно k стандартных.

Решение.

Пространство элементарных исходов Ω создадим из всевозможных наборов по m деталей в каждом из общего множества N деталей. Все эти события будем считать равновероятными. Объем этого пространства равен числу сочетаний из N по m , т.е. числу C_N^m . Остается посчитать число благоприятных исходов. Среди n стандартных деталей выбрать ровно k деталей можно C_n^k способами; при этом остальные $m - k$ деталей должны быть нестандартными. Взять $m - k$ нестандартных деталей среди $N - n$ нестандартных можно C_{N-n}^{m-k} способами. Следовательно, число благоприятных исходов равно $C_{N-n}^{m-k} C_n^k$. Исходная вероятность, согласно формуле классической вероятности, равна:

$$P = \frac{C_{N-n}^{m-k} C_n^k}{C_N^m}.$$

2.3. Геометрическая вероятность

Предположим, что случайно выбирается точка на отрезке $[a, b]$. Поставим задачу об определении вероятности попадания в точку $x_0 \in [a, b]$. Если исходить из равновозможности выбора произвольной точки и классической формулы вероятности для равновозможных исходов, то мы должны положить $p(x_0) = \frac{1}{\infty} = 0$. Тот же результат получим, если исходить из аксиомы непрерывности A_4 : если ω_i — событие, заключающееся в том, что случайно выбранная точка попала в ω_i -ый интервал, а ω_i — набор окрестностей точки x_0 , стягивающихся в точку x_0 , то $\lim_{i \rightarrow \infty} p(\omega_i) = 0$, согласно аксиоме A_4 (подробнее об аксиоме непрерывности в п.3).

Но если вероятность события равна нулю, то само событие "невозможное", т.е. никакая точка не может быть выбрана, что по условию задачи тоже невозможно. Мы пришли к так называемому "парадоксу непрерывности".

Мы должны сделать вывод о том, что классическое определение вероятности нельзя применить к опыту с бесконечным числом равновозможных исходов, т.к. вероятность каждого события равна нулю по классической формуле и любой элементарный исход невозможен. Но на самом деле какое-то событие из пространства элементарных исходов обязательно произойдет.

Выход из парадокса непрерывности следующий. Вместо вероятности "попадания в точку" вводится вероятность "попадания в область ω , содержащую точку", и "относительная величина области" приравнена вероятности попадания в любую точку этой области $f(x_0) = \frac{\mu(\omega)}{\mu(\Omega)}$, где $\mu(\omega)$, $\mu(\Omega)$ — мера (длина, площадь, объем, время и др.) области ω и Ω соответственно.

Такая ситуация возникает, когда пространство элементарных исходов представляет собой непрерывное множество (прямая или ее часть, плоская или пространственная область).

Когда объем пространства элементарных исходов состоит из неограниченного числа событий, вероятностное пространство строится на основе σ -алгебры. Пусть в пространстве Ω введена мера μ (длина, площадь, объем, время и др.) и $\mu(\Omega) < \infty$. Функция вероятности на элементах алгебры может быть введена следующим образом:

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}. \quad (2.3.1)$$

Определение 2.3.1. Вероятность, вычисляемая как отношение мер по формуле (2.3.1), называется **геометрической вероятностью**.

Необходимо отметить, что в качестве алгебры уже нельзя брать максимальную, т.е. множество всевозможных подмножеств множества Ω , поскольку в алгебру не могут войти множества, не имеющие меры множества Ω . Например, кусок прямой

на плоскости имеет меру равную длине этой прямой, но эта же мера равна нулю с точки зрения площади.

♦ **Пример 1 (задача Бюффона.)**³ На плоскость, расчерченную параллельными прямыми, находящимися на расстоянии a друг от друга, случайно брошена игла длины $l < a$. Найти вероятность пересечения иглы с какой-нибудь из параллельных прямых.

Решение.

Пусть событие A — пересечение иглы с какой-нибудь из параллельных прямых.

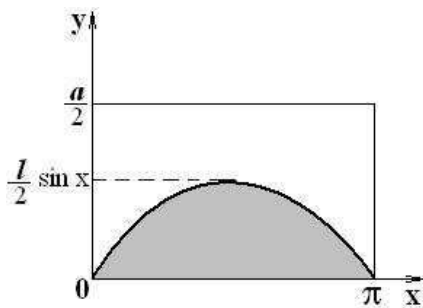


Рис.5

Введем прямоугольную систему координат xOy , и пусть x — угол между иглой и параллельной прямой, y — расстояние от середины иглы до ближайшей прямой. За пространство элементарных событий Ω примем прямоугольник

$\Omega = \{(x, y) : 0 < x < \pi, 0 < y < a/2\}$, $\mu(\Omega) = \pi \cdot \frac{a}{2}$. Интересующее событие A имеет вид: $A = \{(x, y) : 0 < y < l/2 \cdot \sin x, 0 < x < \pi\}$, $\mu(A) = \int_0^\pi \frac{l}{2} \cdot \sin x \, dx$ (рис.5).

³Обычно считают, что именно работы французского естествоиспытателя Ж. Бюффона (1707–1788), в которых он сформулировал задачу о бросании иглы на разграфленную плоскость и предложил ее решение, явились основополагающими в формировании геометрического определения вероятности. Первая его публикация на эту тему относится к 1733 г., когда он сделал доклад в Парижской академии наук под названием "Мемуар об игре франк-карро". Текст второй задачи Бюффона приведен в примере 3.3.1. После Бюффона задачи на геометрические вероятности систематически включались в работы по теории вероятностей. Это исследования П. Лапласа, В.Я. Буныковского, Г. Ламе, Д. Сильвестра, М. Крофтона, Ж. Бертрана и других. Более того, из опыта с "бюффоновой иглой" экспериментально подсчитано число π .

Искомая вероятность равна

$$P = \frac{\mu(A)}{\mu(\Omega)} = \frac{l}{\pi a} \int_0^\pi \sin x \, dx = \frac{2}{\pi} \frac{l}{a}.$$

◇ **Пример 2.** Сигналы от двух независимо работающих анализаторов поступают в сигнализатор в любой момент времени в течение времени T . Если разность между поступлениями этих двух сигналов меньше, чем t , срабатывает сигнализатор. Найти вероятность срабатывания сигнализатора, если известно, что каждый из анализаторов послал сигнал в течение времени T .

Решение.

Введем прямоугольную систему координат xOy и по оси Ox будем откладывать время поступления первого сигнала, а по оси Oy — время поступления второго сигнала. Предупреждение сработает при выполнении неравенства

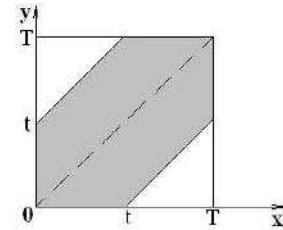


Рис.6

$$|x - y| < t.$$

За пространство элементарных событий Ω примем квадрат $\Omega = \{(x, y) : 0 < x < T, 0 < y < T\}$. Множество точек срабатывания предупреждающего устройства: $B = \{(x, y) : -t < x - y < t\}$ (рис.6).

Искомая вероятность равна

$$P = \frac{\mu(B)}{\mu(\Omega)} = \frac{T^2 - (T - t)^2}{T^2} = \frac{2Tt - t^2}{T^2} = \frac{t(2T - t)}{T^2}.$$

2.4. Статистическая вероятность

Классическое определение вероятности в задачах естественно-научного и технического характера наталкивается на принципиальные трудности. Очевидно, что не любое

событие обладает симметрией, позволяющей сделать заключение о "равновозможности", и совсем не обязательна конечность пространства элементарных исходов.

Представьте себе бесконечный конвейер, по которому движутся детали, среди которых есть бракованные. Как можно определить вероятность появления бракованной детали, если пересчет всех деталей невозможен, а о равновероятных исходах не может быть и речи. В такой ситуации можно лишь сделать "случайную выборку" деталей с конвейера, пересчитать в ней количество бракованных и найти отношение этого количества к объему выборки. Мы здесь следуем обычным канонам теории вероятности. Но можно ли считать полученное число вероятностью? Конечно — нет. Полученное число можно назвать "частотой" появления бракованного изделия именно в этой выборке. Возникает вопрос — это число (частота) характеризует происходящее на конвейере или оно такое же случайное как и весь процесс поиска частоты? Ведь выборка — случайная, количество бракованных деталей — случайное и пр..

Здравый смысл подсказывает, что надо произвести серию таких испытаний, если в результате "частота" окажется "стабильной" (в некотором смысле), то можно считать "частоту" реальной характеристикой процесса и оперировать этой цифрой в сводках, отчетах и научных исследованиях. Но это лишь здравый смысл и интуиция.

Практическую проверку "частоты" можно установить на самом простом опыте — подбрасывании монеты. Вероятность появления "орла", "решки" — $\frac{1}{2}$. Проведите теперь серию подбрасываний с последующим определением частоты в каждой серии. Очень принципиальный вопрос — будет ли частота приближаться к числу — $\frac{1}{2}$? Если нет, то исследование и других процессов на основе "частоты" — пустая забава, если да, то в "частоте" заложен смысл "вероятности", и следовательно, эта характеристика существенна. На проверку уйдет много времени, а справедливым окажется второй ответ ⁴.

⁴Кто и когда впервые проделал опыт с подбрасыванием монеты не из-

Свойство устойчивости частот, многократно проверенное экспериментально, есть одно из наиболее характерных закономерностей, наблюдаемых в случайных явлениях. Математическую формулировку этой закономерности впервые дал Я. Бернулли в теореме, представляющей простейшую форму закона больших чисел. Бернулли доказал, что при неограниченном увеличении числа независимых испытаний с практической достоверностью можно утверждать, что частота события будет сколь угодно мало отличаться от его вероятности в отдельном испытании. Выводы английских ученых Д. Граунта и В. Петти относительно устойчивости частот некоторых событий, в частности госпитальной смертности для Лондона и Парижа, подготовили почву к открытию теоремы, названной позже "Законом больших чисел".

Теперь приведем соответствующее определение.

Пусть эксперимент повторяется n раз. Обозначим через ω_k – исход в k -ом эксперименте, n -кратное повторение эксперимента означает, что определена последовательность $\{\omega_1, \omega_2, \dots, \omega_n\}$. Эту последовательность мы и примем за пространство элементарных исходов Ω^n . Обозначим через μ – число появления события A в n испытаниях. Тогда $\nu_n = \mu/n$ – называется *относительной частотой события A* .

Определение 2.4.1. Если произведена серия из n испытаний, в каждом из которых могло появиться или не появиться событие A , то относительной частотой события A в данной серии испытаний называется отношение числа испытаний μ , в которых появилось событие A , к общему числу всех произведенных испытаний:

$$\nu_n = \mu/n$$

Определение 2.4.2. Под статистической вероятностью события понимают относительную частоту события A при

вестно. Но истории известны результаты испытаний некоторых ученых. Ж. Бюффон (1707–1788) 4040 раз подбрасывал монету, "герб" выпал 2048 раз. К. Пирсон (1857–1936) подбросил ее 24000 раз, "герб" выпал 12012 раз.

большом числе испытаний n :

$$P(A) = \lim_{n \rightarrow \infty} \nu_n.$$

Таким образом, на статистическую вероятность можно смотреть как на вероятность в n исходах, поэтому на это определение переносятся основные свойства вероятности (с равновозможными исходами):

1. $P(\Omega) = 1$;
2. $P(\emptyset) = 0$;
3. Если $C = A_1 + A_2 + \dots + A_n$, где A_1, A_2, \dots, A_n — несовместны, то $P(C) = P(A_1) + P(A_2) + \dots + P(A_n)$.

§ 3. Аксиоматика теории вероятностей

В пункте 2.1 при рассмотрении пространств элементарных равновозможных исходов Ω были получены семь свойств классического определения вероятности. Интерес вызывает вопрос о том, какие из этих свойств можно доказать для произвольного (в том числе с бесконечным объемом) пространства элементарных исходов, какие свойства следует считать аксиомами.

Аксиоматическое построение теории вероятностей⁵ берет начало от анализа классического, геометрического и статистического определений и призвано преодолеть недостатки каждого из них.

⁵Впервые аксиоматический подход, основанный на качественном сравнении случайных событий по их меньшей или большей вероятности, был предложен в 1917 году С.Н. Бернштейном (1880-1968 гг.). Все основные положения теории вероятностей выводились им дедуктивным путем из основных трех аксиом — аксиомы сравнения вероятностей, аксиомы о несовместимых событиях и аксиомы о совмещении событий. В 1933 году вышла книга А.Н. Колмогорова "Основные понятия теории вероятностей в которой была предложена аксиоматика, получившая всеобщее признание и позволившая охватить не только все классические разделы теории вероятностей, но и дать строгую основу для развития новых разделов.

Определение 3.1. Пусть Ω — пространство элементарных исходов, \mathcal{U} — какая-либо алгебра, построенная из подмножеств (событий) множества Ω . **Вероятностью** будем называть функцию $P = P(A)$, определенную на элементах алгебры $A \in \mathcal{U}$, для которой выполнены следующие аксиомы теории вероятностей:

- A1. $P(A) \geq 0 \quad \forall A \in \mathcal{U}$ (неотрицательность P);
- A2. $P(\Omega) = 1$ (нормированность P);
- A3. $AB = \emptyset, \implies P(A + B) = P(A) + P(B)$ (аддитивность P);
- A4. Для любой последовательности $A_j \in \mathcal{U}, j = 1, 2, \dots$ включающих друг друга событий $A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \supset \dots$, для которых

$$\bigcap_{n=1}^{\infty} A_n = \emptyset$$

выполняется равенство

$$\lim_{n \rightarrow \infty} P(A_n) = 0.$$

Из аксиом выведем некоторые важные свойства.

Свойство 1). $P(\bar{A}) = 1 - P(A)$.

Следует из A3 и A2, поскольку $A + \bar{A} = \Omega, \quad A \cdot \bar{A} = \emptyset$.

Свойство 2). Если $A_i \cdot A_j = \emptyset, \forall i \neq j$, то имеет место счетная аддитивность:

$$P\left(\sum_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k).$$

Следует из A3, например, для трех слагаемых

$$P(A_1 + A_2 + A_3) = P(A_1) + P(A_2 + A_3) = P(A_1) + P(A_2) + P(A_3).$$

Свойство 3). $P(\emptyset) = 0$.

Следует из A2 и свойства 1.

Свойство 4). Если $A \subset B$, то $P(B - A) = P(B) - P(A)$.

Для доказательства представим событие B в виде $B = A + (\bar{B} \cap A)$.
Т.к. $A \cap (B - A) = \emptyset$, то по АЗ: $P(B) = P(A) + P(B - A)$.

Свойство 5). Если $A \subset B$, то $P(A) \leq P(B)$.

Следует из свойства 4 и A_1

Свойство 6). $0 \leq P(A) \leq 1$.

С очевидностью вытекает из свойства 5.

Свойство 7). $P(A + B) = P(A) + P(B) - P(AB)$.

Для доказательства представим события A и $A + B$ в виде суммы непересекающихся событий:

$$A = A\bar{B} + AB, \quad A\bar{B} \cap AB = \emptyset \xrightarrow{A3} P(A) = P(A\bar{B}) + P(AB).$$

$$A + B = A\bar{B} + B, \quad A\bar{B} \cap B = \emptyset \xrightarrow{A3} P(A + B) = P(A\bar{B}) + P(B).$$

Из первого получаем:

$$P(A\bar{B}) = P(A) - P(AB)$$

и, подставляя во второе, находим требуемое.

Свойство 8). Для любых событий A_1, A_2, \dots, A_n алгебры \mathcal{U} выполняется неравенство

$$P\left(\sum_{k=1}^n A_k\right) \leq \sum_{k=1}^n P(A_k).$$

Следует из свойства 7, так как $P(A \cdot B) \geq 0$, то $P(A + B) \leq P(A) + P(B)$. Индукцией индексу суммирования выводим требуемое.

Замечание 1. Свойство 7 носит название "теоремы сложения" для произвольных событий A и B (ср. с примером 1 пункта 2.1).

Замечание 2. Аксиома 4 позволяет интерпретировать себя так: достаточно "мелкому" событию отвечает достаточно малая вероятность. Поэтому эта аксиома называется "аксиомой непрерывности".

Отметим также следующее. А4 рассматривается только на элементах алгебры, построенной на основе пространства элементарных событий Ω с бесконечным объемом (иногда говорят — с

бесконечной размерностью) $n = \infty$, и может быть заменена на **расширенную аксиому сложения**: если события A_n в последовательности A_1, A_2, \dots попарно несовместны, то

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Это вытекает из следующей теоремы.

Теорема 3.1. *Расширенная аксиома сложения равносильна аксиоме непрерывности.*

Доказательство.

1. Из расширенной аксиомы сложения следует аксиома непрерывности. Действительно, пусть $A_1, A_2, \dots, A_n, \dots$ таковы, что $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$ при $\forall n \geq 1$ и $\bigcap_{n=1}^{\infty} A_n = \emptyset$. Очевидно, что $A_n = \sum_{k=n}^{\infty} A_k \cdot \bar{A}_{k+1} + \bigcap_{k=n}^{\infty} A_k$. Так как события, стоящие в сумме, попарно несовместны, то по расширенной аксиоме сложения

$$P(A_n) = \sum_{k=n}^{\infty} P(A_k \cdot \bar{A}_{k+1}) + P\left(\bigcap_{k=n}^{\infty} A_k\right).$$

Последнее слагаемое равно нулю по условию, следовательно

$$P(A_n) = \sum_{k=n}^{\infty} P(A_k \cdot \bar{A}_{k+1}),$$

то есть $P(A_n)$ есть остаток сходящегося ряда

$$\sum_{k=1}^{\infty} P(A_k \cdot \bar{A}_{k+1}) = P(A_1),$$

поэтому

$$\lim_{n \rightarrow \infty} P(A_n) = 0.$$

2. Из аксиомы непрерывности следует аксиома сложения. Пусть $A_1, A_2, \dots, A_n, \dots$ попарно несовместны и $A = A_1 + A_2 + \dots + A_n + \dots$. Положим

$$B_n = \sum_{k=n}^{\infty} A_k,$$

тогда $B_{n+1} \subset B_n$. Если событие B_n наступило, то наступило какое-нибудь из событий A_i ($i \geq n$) и, значит, в силу попарной несовместности событий A_k , события A_{i+1}, A_{i+2}, \dots не наступили. Таким образом, события B_{i+1}, B_{i+2}, \dots — невозможны, и следовательно, невозможно событие $\bigcap_{k=n}^{\infty} B_k$. По аксиоме непрерывности

$$\lim_{n \rightarrow \infty} P(B_n) = 0.$$

Так как $A = A_1 + A_2 + \dots + A_n + B_{n+1}$, то по аксиоме аддитивности

$$\begin{aligned} P(A) &= P(A_1) + P(A_2) + \dots + P(A_n) + P(B_{n+1}) = \lim_{n \rightarrow \infty} \sum_{k=1}^n P(A_k) = \\ &= \sum_{k=1}^{\infty} P(A_k). \end{aligned}$$

Доказательство закончено.

Обратите внимание, что числовая функция вероятности при аксиоматическом подходе вводится, вообще говоря, произвольно. Исследователь, занимающийся решением практической задачи, должен отдавать себе отчет в том, что:

- 1) для одного и того же опыта можно построить разные пространства элементарных событий Ω ;
- 2) на пространстве элементарных событий можно построить разные алгебры;
- 3) числовая функция вероятности $P(A)$ на элементах одной и той же алгебры \mathcal{U} может быть введена различными способами.

Сделав эти три выбора, т.е. определив тройку (Ω, \mathcal{U}, P) , вы построите "вероятностное пространство" испытания или, как часто говорят, "вероятностную модель". Насколько удачно введено вероятностное пространство, нужно судить по согласованности теоретических выводов с данными статистического опыта.

Приведем соответствующие примеры.

- ◇ 1) В "монетку" играют три человека. Бросаются две монеты, первый выигрывает если выпало два "орла", второй — две "решки", третий — монеты упали разными сторонами. Предположение, что монеты могут упасть разными или одинаковыми

сторонами, приводит к пространству элементарных исходов⁶

$$\Omega = \{(O, R); (O, O); ((R, R))\}.$$

Если считать эти события равновероятными, то вероятность каждого из элементарных исходов следует считать равной $1/3$.

Однако, зафиксировав каким-либо образом монеты, мы можем описать пространство элементарных исходов четырьмя исходами

$$\Omega = \{(O, R); (R, O); (O, O); ((R, R))\}.$$

Если посчитать эти события равновероятными, то вероятность выигрыша первым игроком $1/2$, а не $1/3$, как это было в первом предположении. Какая из этих вероятностных моделей игры более "справедливая", может указать опыт, который читатель может сделать самостоятельно (потратив на это немало времени), а также может обратиться к исследованиям Паскаля парадокса де Мере, приведенным во введении.

Подбор нужной вероятностной модели осуществляется с помощью соответствующего эксперимента. В рассмотренном выше примере по сути мы должны выяснить, ведут ли себя монеты как одинаковые (модель Ω_1) или как разные (модель Ω_2 , где пары упорядочены).

◇ 2) $\Omega = \{0; 1\}$, $\mathcal{U}_1 = \{(\Omega; (0, 1))\}$, $\mathcal{U}_2 = \{(\Omega; (0); (1); (0, 1))\}$. Предполагается, что элементы $(0, 1)$ и $(1, 0)$ совпадают.

◇ 3) $\Omega = \{A_1, \dots, A_n\}$. Вероятность вводится с помощью произвольного набора неотрицательных чисел p_j ($p_j \leq 1$), удовлетворяющих условию $\sum_{j=1}^n p_j = 1$. Ясно, что таких наборов —

⁶Подобную ошибку в подсчете вероятностей при игре в орлянку допустил французский математик и философ Ж.Л. Даламбер (1717–1783) в статье "Герб и решка" (1754), которая стала печально знаменитой. Но это была не единственная ошибка в исследованиях ученого. Много неточностей было допущено в приложениях теории вероятностей к статистике народонаселения, в рассуждениях о "физической" и "математической" вероятностях и др.. Причиной тому служило по-видимому общее критическое отношение Ж.Л. Даламбера к теории вероятностей.

бесчисленное множество. Если события A_j равновозможны, то естественно считать $P(A_j) = 1/n$.

§ 4. Условные вероятности. Независимость

4.1. Условная вероятность и теоремы умножения

При решении задач о нахождении вероятности часто приходится сталкиваться с ситуацией, когда кроме исходных данных о событии появляется информация о событиях, которые уже произошли, и тогда происходит изменение вероятностного пространства. Например, если студент знает 10 из 21 вопроса программы, то пространство элементарных исходов состоит из 21 вопроса, алгебра предполагается минимальной $\mathcal{U} = (\Omega, \emptyset)$ (т.е. студенту зададут только один вопрос)⁷, вероятность "счастливого" вопроса — $\frac{10}{21}$. Но представьте себе, что два вопроса уже заданы и оба из счастливой десятки. Пространство элементарных исходов автоматически сужается до 19 вопросов, а вероятность счастливого вопроса уже — $\frac{8}{19}$. Этот пример приводит нас к необходимости рассматривать события *безусловные* — известно только вероятностное пространство (Ω, \mathcal{U}, P) и *условные* — есть дополнительная информация, позволяющая сузить пространство элементарных исходов, что изменит и вероятностное пространство. Соответственно появляются термины *безусловная вероятность* и *условная вероятность*.

Условные вероятности обозначаются $P(A/B)$ — вероятность события A при условии, что произошло событие B . Пример нахождения условной вероятности для вероятностного пространства (Ω, \mathcal{U}, P) с равновозможными элементарными исходами $\omega_i \in \Omega$ и вероятностью $P(\omega_i) = 1/n$ не представляет труда и опреде-

⁷Конечно, в действительности студенту могут задать не один вопрос и даже "прочитать мораль". Но в рамках данного вероятностного пространства — это все пустое множество \emptyset .

ляется по формуле (см. пункт (2.1), пример 2):

$$P(A/B) = \frac{P(AB)}{P(B)}.$$

Аналогичная формула получается для вероятностного пространства с геометрической вероятностью:

$$P(A/B) = \frac{\mu(AB)}{\mu(B)} = \frac{\mu(AB)/\mu(\Omega)}{\mu(B)/\mu(\Omega)} = \frac{P(AB)}{P(B)}.$$

Здесь нужно обратить внимание на следующий факт: вероятности в полученных равенствах взяты из первоначального вероятностного пространства (Ω, \mathcal{U}, P) , что делает формулы удобными не только для практического применения. Это дает основание для введения следующего определения *условной вероятности* в произвольном вероятностном пространстве.

Определение 4.1.1. Пусть (Ω, \mathcal{U}, P) — произвольное вероятностное пространство. Если $A, B \in \mathcal{U}$, $P(B) > 0$, то условная вероятность события A при условии, что произошло событие B , определяется формулой

$$P(A/B) = \frac{P(AB)}{P(B)}. \quad (4.1.1)$$

В правой части равенства (4.1.1.) символ P понимается как вероятность в рассматриваемом вероятностном пространстве, а слева, как вероятность в новом вероятностном пространстве, определенном условием, что событие B произошло.

Разумеется, при этом функция вероятности, определенная на условных событиях A/B по формуле (4.1.1), должна удовлетворять аксиомам теории вероятности $A1 - A4$. Проверьте это самостоятельно, зафиксировав некоторое событие B с $P(B) > 0$.

Из определения (4.1.1) с очевидностью вытекает

Теорема умножения 1.

$$P(AB) = \begin{cases} P(A) P(B/A) \\ P(B) P(A/B) \end{cases} \quad (4.1.2)$$

Общий вариант этой теоремы.

Теорема умножения 2.

$$P\left(\prod_{i=1}^n A_i\right) = P(A_1)P(A_2/A_1)P(A_3/A_1A_2) \dots P(A_n/A_1A_2 \dots A_{n-1}). \quad (4.1.3)$$

Доказательство. При $n = 2$ равенство (4.1.2) вытекает непосредственно из (4.1.1). Далее по индукции. Предположим, что (4.1.3) справедливо для $n - 1$ сомножителей:

$$P\left(\prod_{i=1}^{n-1} A_i\right) = P(A_1)P(A_2/A_1) \dots P(A_{n-1}/A_1A_2 \dots A_{n-2}). \quad (4.1.4)$$

Тогда n сомножителей $A_1 \dots A_n$ рассмотрим как произведение двух событий: A_n и $A_1A_2 \dots A_{n-1}$, к которым применима формула (4.1.2) для двух событий:

$$P\left(\prod_{i=1}^n A_i\right) = P\left(\prod_{i=1}^{n-1} A_i\right) P(A_n/A_1 \dots A_{n-1}).$$

Подстановка в это выражение (4.1.4) и даст формулу (4.1.3).

Доказательство закончено⁸.

4.2. Независимость событий

Определение 4.2.1. События A и B называются независимыми, если

$$P(AB) = P(A)P(B).$$

◇ **Пример 1.** Опыт состоит в двукратном подбрасывании монеты. Пусть событие A означает выпадение "орла" в первом

⁸Теорема умножения событий имела достаточно длинную историю. Многие ученые, в том числе Я. Бернулли, П.Р. Монмор и другие, рассматривали ее на частных примерах при подсчете числа шансов, благоприятствующих наступлению двух и более событий. Впервые формулировку теоремы умножения и введение условной вероятности удалось осуществить английскому математику А. Муавру в 1718 году в первом издании "Доктрины шансов"

подбрасывании, событие В — выпадение "решки" во втором подбрасывании. Зависимы ли события А и В?

Решение. Пространство элементарных событий имеет вид: $\Omega = \{OO, OR, RO, RR\}$.

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{2}, \quad P(AB) = \frac{1}{4}.$$

Так как $P(AB) = P(A)P(B)$, то события А и В независимы.

Отметим некоторые свойства независимых событий:

1. Пусть $P(B) > 0$, то события А и В независимы тогда и только тогда, когда

$$P(A/B) = P(A).$$

Если $P(A) > 0$, то события А и В независимы тогда и только тогда, когда

$$P(B/A) = P(B).$$

2. Два события, одно из которых является невозможным событием, независимы.

3. Если события А и В независимы, то независимы события А и \bar{B} , \bar{A} и В, \bar{A} и \bar{B} .

Доказательство.

Свойство 1 с очевидностью вытекает из определения условной вероятности. Таким образом, понятие независимости событий А и В симметрично.

Свойство 2 очевидно. Например, пусть

$$A = \emptyset, \Rightarrow A \cdot B = \emptyset, \Rightarrow P(AB) = 0.$$

С другой стороны

$$P(A) \cdot P(B) = 0 \cdot P(B) = 0.$$

А значит, $P(AB) = P(A) \cdot P(B)$, т.е. А и В независимы.

Для доказательства свойства 3 представим событие А в виде $A = (AB) + (A\bar{B})$. События АВ и $A\bar{B}$ несовместны, следовательно, по третьей аксиоме $P(A) = P(AB) + P(A\bar{B})$. Из последнего равенства и из независимости событий А и В вытекает

$$P(A\bar{B}) = P(A) - P(AB) = P(A) - P(A)P(B) = P(A)(1 - P(B)) =$$

$$= P(A)P(\overline{B}).$$

Таким образом, независимость событий A и \overline{B} доказана. Аналогично доказываются остальные утверждения свойства 3.

Определение 4.2.2. События A_1, A_2, \dots, A_n называются независимыми в совокупности, если для любых

$$1 \leq i_1 < i_2 < \dots < i_m \leq n, \quad 1 \leq m \leq n$$

выполняются равенства

$$P(A_{i_1} A_{i_2} \dots A_{i_m}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_m}).$$

Математическое определение независимости, как правило, согласуется с обычным представлением о независимости. Например, если в опыте с бросанием двух игральных костей выпадение определенной грани на одной кости и выпадение этого же количества очков на другой кости независимы в общепринятом смысле, то эти события независимы и в математическом смысле. Но подобное согласование наблюдается не всегда. Отметим, что независимость нескольких событий называется *независимостью событий в совокупности*. Рассмотрим пример показывающий, что независимость событий в совокупности — более сильное свойство, чем попарная их независимость.

♦ **Пример 2**⁹. Положим, что в ящике находятся 4 одинаковых билетика, так что вероятность вынуть каждый билетик одна и та же. Пусть номера этих билетов будут соответственно 112, 121, 211 и 222. Событие A_1 означает, что цифра 1 стоит на месте сотен, событие A_2 — цифра 1 стоит на месте десятков, событие A_3 — цифра 1 стоит на месте единиц в вынутом номере. Докажите попарную независимость событий A_i , $i = 1, 2, 3$, и что все эти три события не являются независимыми в совокупности.

Решение.

⁹Это классический пример С.Н. Бернштейна, опубликованный в 1927 году в учебном пособии для высшей школы "Теория вероятностей", в котором автор доказывает, что "попарная независимость между тремя событиями отнюдь не означает, что эти события совершенно независимы".

Очевидно, что $P(A_1) = P(A_2) = P(A_3) = 2/4 = 1/2$. Следовательно, $P(A_1)P(A_2) = P(A_1)P(A_3) = P(A_2)P(A_3) = 1/4$. И $P(A_1A_2) = P(A_1A_3) = P(A_2A_3) = 1/4$. Таким образом,

$$P(A_1A_2) = P(A_1)P(A_2), \quad P(A_1A_3) = P(A_1)P(A_3),$$

$$P(A_2A_3) = P(A_2)P(A_3),$$

что означает попарную независимость событий A_i . Однако

$$P(A_1A_2A_3) = P(\emptyset) \neq 1/8 = P(A_1)P(A_2)P(A_3),$$

и следовательно, эти три события A_1, A_2, A_3 не являются независимыми в совокупности.

4.3. Формула полной вероятности. Формулы Байеса

♦ **Пример 1.** На склад поступают изделия трех заводов. Тогда событие $A = \{\text{взятое со склада изделие — брак}\}$ произойдет вместе с событием — $H_i = \{\text{изделие изготовлено конкретным заводом}\}$ $i = 1, 2, 3$. Фиксация i , разумеется, лишь предположение (гипотеза). Таким образом, имеем три гипотезы, одна из которых всегда произойдет, т.к. по условию изделия поступили только с данной группы заводов. Тогда событие $A = \{\text{бракованное изделие изготовлено или первым заводом, или вторым заводом, или третьим заводом}\}$:

$$A = AH_1 + AH_2 + AH_3.$$

Теорема. Пусть задано вероятностное пространство $\{\Omega, \mathcal{U}, P\}$ и для события $A \in \mathcal{U}$ выделен набор, так называемых, "гипотез":

$$\{H_i\}_{i=\overline{1,n}}, \quad H_iH_j = \emptyset, \quad A \subset H_1 + H_2 + \dots + H_n.$$

Тогда имеют место следующие формулы

$$P(A) = \sum_{k=1}^n P(H_k) P(A/H_k). \quad (4.3.1)$$

$$P(H_k/A) = \frac{P(H_k) P(A/H_k)}{\sum_{i=1}^n P(H_i) P(A/H_i)}. \quad (4.3.2)$$

Первая из них называется "формулой полной вероятности", вторая — "формулы Байеса"¹⁰.

Доказательство.

Так же, как и в приведенном выше примере, событие A можно представить в виде следующей суммы попарно несовместных событий (рис.7)

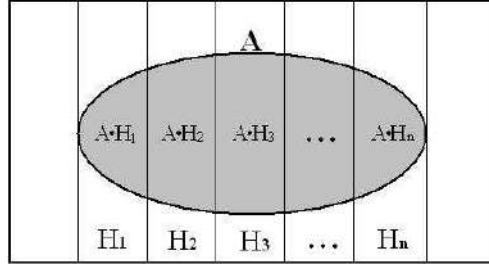


Рис.7

$$A = AH_1 + AH_2 + \dots + AH_n.$$

Отсюда воспользовавшись аксиомой А3 и формулой (4.1.3), получим формулу (4.3.1). Доказательство закончено.

Теперь, согласно определению "условной вероятности"

$$P(H_k/A) = \frac{P(AH_k)}{P(A)},$$

заменяв здесь числитель по формуле (4.1.2), а знаменатель по формуле полной вероятности (4.3.1), получим *формулу Байеса*.

♦ **Пример 2.** Из урны, содержащей M белых шаров и $N - M$ черных, утерян один шар неизвестного цвета. Какова вероятность извлечь из урны белый шар?

¹⁰Используя результаты теоремы умножения, полученные Муавром, Т. Байес (1702–1761) вывел несколько следствий для вычисления вероятности $P(B/A)$ по вероятностям $P(AB)$ и $P(A)$. Словесную формулировку формул Байеса и полной вероятности впервые дал французский астроном, математик и физик П.С. Лаплас в работе "Опыт философии теории вероятностей" в виде основных принципов теории вероятностей (6 и 7 принципы соответственно). Название "формулы Байеса" появилось в середине XIX века.

Р е ш е н и е. Возможны две гипотезы: $H_1 = \{\text{утерян белый шар}\}$ и $H_2 = \{\text{утерян черный шар}\}$. Обозначим через $A = \{\text{шар, извлеченный из оставшихся шаров, оказался белым}\}$. Имеем: $P(H_1) = \frac{M}{N}$, $P(H_2) = \frac{N-M}{N}$. По формуле полной вероятности

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) = \frac{M}{N} \frac{M-1}{N-1} + \\ + \frac{N-M}{N} \frac{M}{N-1} = \frac{M}{N}.$$

Отметим, что вероятность извлечения белого шара из урны до утери шара тоже равнялась $\frac{M}{N}$.

§ 5. Последовательность испытаний

Всякий знает, что последовательности случайных событий могут так сильно влиять на настоящее, что перестают казаться случайными. Поэтому большую значимость приобретает изучение зависимости от истории (возможно, длинной) предыдущей цепи событий. Разумеется, самые общие последовательности, наверное, не поддадутся полному математическому описанию, поскольку точно даже не известно, случайны ли они. Но некоторые из этих последовательностей случайных событий допускают простое и, стало быть, красивое описание. Простой пример. Представьте, что некоторый опыт может закончиться N исходами (бросили кубик — 6 исходов), и производится n опытов. Элементарное событие можно описать парой натуральных чисел i_j , где i — номер опыта (меняется от 1 до n), а j — номер исхода (меняется от 1 до N).

Тогда множеством элементарных событий является множество:

$$\Omega_n = \{(i_1, i_2, \dots, i_n), \quad i = 1, 2, \dots, N\}.$$

Элементарное событие $\omega = i_1 i_2 \dots i_n$ интерпретируется как цепочка исходов в n последовательных испытаний, каждое из

которых имеет N несовместных исходов. По теореме умножения имеем:

$$p(\omega) = p(i_1)p(i_2/i_1) \dots p(i_n/i_1 \dots i_{n-1}),$$

где $p(i_s/i_1 \dots i_{s-1}) \geq 0$, и $\sum_{i_s=1}^N p(i_s/i_1 \dots i_{s-1}) = 1$,

$$s = 1, \dots, n.$$

Тогда на множестве Ω_n однозначно определяется вероятность:

$$P(A) = \sum_{\omega \in A} p(\omega), \quad A \subset \Omega_n.$$

Можно проверить, что $P(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$.

Ниже мы познакомимся с двумя наиболее простыми последовательностями испытаний, одно из которых строится так, чтобы вероятность следующего события зависела только от предыдущего (цепь Маркова), а другое — так, чтобы вероятность следующего никак не зависела от всей предыдущей последовательности событий (схема испытаний Бернулли).

5.1. Понятие цепи Маркова

Рассмотрим последовательность испытаний, в каждом из которых происходит или не происходит некоторое событие A . И пусть A_i обозначает, что событие A произошло в результате i -го испытания.

Определение 5.1.1. *Цепью Маркова называется такая последовательность испытаний, что условная вероятность $P(A_k/A_1A_2 \dots A_{k-1})$ события A_k зависит только от предыдущего испытания A_{k-1} и не зависит от A_1, A_2, \dots, A_{k-2} .*

В качестве примера такой последовательности рассмотрим следующий процесс. Пусть имеются три урны, содержащие белые и черные шары:

- в первой — 2 белых, 3 черных;
- во второй — 2 белых, 2 черных;
- в третьей — 3 белых, 1 черный.

Из первой урны случайным образом выбирается шар и перекладывается во вторую, затем из второй урны случайным образом выбирается шар и перекладывается в третью, а затем — из третьей в первую. Предположим, что нас интересует, какой состав шаров в первой урне после перекладываний наиболее вероятен.

Ясно, что исходов может быть только три:

B_1 — в первой урне оказался один белый шар;

B_2 — в первой урне оказалось два белых шара;

B_3 — в первой урне оказалось три белых шара.

Через A_i (соответственно через \bar{A}_i) обозначим событие, состоящее в том, что на i -ом шаге перекладывается белый (черный) шар.

Пространством элементарных исходов является следующий набор несовместных событий

$$\Omega = \{A_1 A_2 A_3, \bar{A}_1 A_2 A_3, A_1 \bar{A}_2 A_3, A_1 A_2 \bar{A}_3, \bar{A}_1 \bar{A}_2 \bar{A}_3, A_1 \bar{A}_2 \bar{A}_3, \bar{A}_1 A_2 \bar{A}_3, \bar{A}_1 \bar{A}_2 A_3\}.$$

Из них легко определить события, составляющие B_1 , B_2 и B_3 . Как видим

$$B_1 = A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 \bar{A}_3;$$

$$B_2 = A_1 A_2 A_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 \bar{A}_3 + \bar{A}_1 \bar{A}_2 \bar{A}_3;$$

$$B_3 = \bar{A}_1 A_2 A_3 + \bar{A}_1 \bar{A}_2 A_3.$$

Отсюда, учитывая несовместность слагаемых событий по третьей аксиоме вероятности, имеем:

$$P(B_1) = P(A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 \bar{A}_3) = P(A_1 A_2 \bar{A}_3) + P(A_1 \bar{A}_2 \bar{A}_3);$$

$$P(B_2) = P(A_1 A_2 A_3) + P(A_1 \bar{A}_2 A_3) + P(\bar{A}_1 A_2 \bar{A}_3) + P(\bar{A}_1 \bar{A}_2 \bar{A}_3);$$

$$P(B_3) = P(\bar{A}_1 A_2 A_3) + P(\bar{A}_1 \bar{A}_2 A_3).$$

Вероятности произведений вычисляются по теореме умножения. Например,

$$P(A_1 A_2 A_3) = P(A_1) P(A_2/A_1) P(A_3/A_1 A_2).$$

Вычисление сомножителей в этом равенстве происходит так. В первой урне 5 шаров, среди которых 2 белых, следовательно, $P(A_1) = \frac{2}{5}$. Если A_1 произошло, то во второй урне оказалось 5 шаров, среди которых 3 белых, поэтому $P(A_2/A_1) = \frac{3}{5}$. Если произошло A_2 , то в третьей урне оказалось 5 шаров, среди которых 4 белых, т.е. $P(A_3/A_1 A_2) = \frac{4}{5}$. В последнем случае уместно заметить, что эта вероятность не изменится, если A_1 заменить на $\overline{A_1}$. Т.е. вероятность того, что из третьей урны извлечен белый шар, зависит от того, какой шар переложен из второй урны, но не зависит от исхода первого испытания. По определению эта последовательность испытаний, начавшаяся извлечением шара из первой урны и закончившаяся перекладыванием шара из третьей в первую урну, — *цепь Маркова*.

Возвращаясь к нашей задаче о наиболее вероятном составе шаров в первой урне, вычислим

$$P(A_1 A_2 A_3) = P(A_1) P(A_2/A_1) P(A_3/A_1 A_2) = \frac{2}{5} \frac{3}{5} \frac{4}{5} = \frac{24}{125}.$$

Аналогично

$$P(\overline{A_1} A_2 A_3) = \frac{24}{125}, \quad P(A_1 \overline{A_2} A_3) = \frac{12}{125}, \quad P(A_1 A_2 \overline{A_3}) = \frac{6}{125},$$

$$P(\overline{A_1} \overline{A_2} \overline{A_3}) = \frac{18}{125}, \quad P(A_1 \overline{A_2} \overline{A_3}) = \frac{8}{125}, \quad P(\overline{A_1} A_2 \overline{A_3}) = \frac{6}{125},$$

$$P(\overline{A_1} \overline{A_2} A_3) = \frac{27}{125}.$$

Проверьте, что $P(\Omega) = 1$. Решение поставленной задачи легко вытекает из сравнения вероятностей

$$P(B_1) = \frac{6}{125} + \frac{8}{125} = \frac{14}{125}.$$

$$P(B_2) = \frac{24}{125} + \frac{12}{125} + \frac{6}{125} + \frac{18}{125} = \frac{60}{125}.$$

$$P(B_3) = \frac{24}{125} + \frac{27}{125} = \frac{51}{125}.$$

Убедитесь, что $P(B_1 + B_2 + B_3) = 1$.

Общий случай. Пусть пространством элементарных событий является множество

$$\Omega_n = \{(i_1 i_2 \dots i_n), \ i_k = 1, 2, \dots, N; \ k = 1, 2, \dots, n\}.$$

Элементарное событие интерпретируется как цепочка исходов в n последовательных испытаниях, каждое из которых имеет N несовместных исходов: $1, 2, \dots, N$. Причем

$$p(\omega) = p(i_1)p(i_2/i_1) \dots p(i_n/i_1 \dots i_{n-1}).$$

Если эта последовательность испытаний — цепь Маркова, то

$$p(\omega) = p(i_1)p(i_2/i_1)p(i_3/i_2) \dots p(i_n/i_{n-1}).$$

При этом

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

а на подмножествах $A \in \Omega$ однозначно определяется вероятность

$$P(A) = \sum_{\omega \in A} p(\omega), \quad A \subset \Omega.$$

5.2. Схема испытаний Бернулли

Это частный случай цепи Маркова, когда условная вероятность события A_i (напомним, что A_i — событие A произошло в i -ом испытании) совпадает с вероятностью события A_i : $P(A_i/A_1 A_2 \dots A_{i-1}) = P(A_i)$ и всего два исхода в каждом испытании ($N = 2$): или A , или \bar{A} ¹¹.

¹¹ Дальнейшие рассуждения не связаны с цепями Маркова.

Определение 5.2.1. Последовательность независимых испытаний, в каждом из которых событие A появляется с вероятностью $P(A) = p$ и не появляется с вероятностью

$$P(\bar{A}) = 1 - p = q,$$

называется схемой (испытаний) Бернулли.

Простейшим примером схемы Бернулли является многократное подбрасывание монетки. Действительно, при этом вероятность выпадения "орла" или "решки" все время постоянна, равна $1/2$, т.е. не зависит от номера испытания.

Наиболее распространенные задачи:

определить вероятность, что в n независимых испытаниях событие A появится m раз (не менее m раз, не более m раз);

определить наиболее вероятное число появления события A в n испытаниях схемы Бернулли.

Первая задача решается применением **формулы Бернулли**:

$$P_n(m) = C_n^m p^m q^{n-m}, \quad C_n^m = \frac{n!}{m!(n-m)!}, \quad (5.2.1)$$

(задача "не менее m раз" или "не более m раз" так же решается применением (5.2.1), в суммах соответственно $\sum_{k=m}^n P_n(k)$ или

$$\sum_{k=1}^m P_n(k),$$

а вторая — нахождением наиболее вероятного числа k , удовлетворяющего двойному неравенству

$$np - q \leq k \leq np + p, \quad (5.2.2)$$

причем из того, что отрезок $[np - q, np + p]$ имеет длину, равную единице, следует, что:

а) если число $(np - q)$ — дробное, то существует только одно наиболее вероятное число;

б) если число $(np - q)$ — целое, то существует два наиболее вероятных числа, а именно $k = np - q$ и $k + 1 = np + p$;

с) если число np — целое, то наивероятнейшее число равно np .

Доказательство формулы Бернулли (5.2.1). Обозначим через A_i событие, состоящее в том, что в i -ом испытании произошло событие A , и \bar{A}_i — событие, состоящее в том, что в i -ом испытании не произошло событие A .

Для n испытаний пространство элементарных исходов Ω состоит из произведений типа $\prod_{i,j} \bar{A}_i A_j$, где количество сомножителей постоянно и равно n . Его подпространство " m из n " состоит из произведений, где количество событий A_i (т.е. без черты) равно m , а количество отрицаний \bar{A}_j равно $n-m$. Например, подпространство " 2 из 3 -х" состоит из событий

$$\{(A_1, A_2, \bar{A}_3); (A_1, A_3, \bar{A}_2); (A_2, A_3, \bar{A}_1)\},$$

" 1 из 3 -х" — из событий

$$\{(A_1, \bar{A}_2, \bar{A}_3); (A_2, \bar{A}_1, \bar{A}_3); (A_3, \bar{A}_1, \bar{A}_2)\}.$$

Ясно, что элементы подобных подпространств — несовместны¹², а по условию схемы Бернулли события A_i независимы. Из теорем сложения (на самом деле из аксиомы $A3$) и умножения вытекает

$$\begin{aligned} P_n(m) &= P\left(\sum \prod A_{i_k} \bar{A}_{j_l}\right) = \sum \prod_{k=1}^m P(A_{i_k}) \prod_{l=m}^n P(\bar{A}_{j_l}) = \\ &= \sum p^m q^{n-m}. \end{aligned}$$

Количество же слагаемых равно числу сочетаний из n элементов по m . Отсюда следует формула (5.2.1).

Доказательство (5.2.1) закончено.

Доказательство (5.2.2). Учитывая, что наивероятнейшее число k появления события A в n независимых испытаниях —

¹²В самом деле, могут ли произойти вместе события $(A_1, \bar{A}_2, \bar{A}_3)$ и $(A_2, \bar{A}_1, \bar{A}_3)$, если в первом случае A произошло в первом испытании, а во втором случае — во втором испытании.

такое число, для которого вероятность $P_n(k)$ наибольшая среди других вероятностей $P_n(m)$, можем записать неравенства

$$P_n(k-1) \leq P_n(k), \quad P_n(k+1) \leq P_n(k). \quad (5.2.3)$$

Рассмотрим первое из них. Заменяя здесь вероятности по формуле Бернулли, получим:

$$\begin{aligned} C_n^{k-1} p^{k-1} q^{n-k+1} &\leq C_n^k p^k q^{n-k}, \implies \frac{C_n^{k-1}}{C_n^k} \leq \frac{p}{q}, \implies \\ \implies \frac{n!}{(k-1)!(n-k+1)!} \frac{k!(n-k)!}{n!} &\leq \frac{p}{q}, \\ \implies \frac{k}{n-k+1} &\leq \frac{p}{q}, \implies \\ \implies kq &\leq np - kp + p, \implies k(p+q) \leq np + p. \end{aligned}$$

Но $p+q=1$, и мы получаем правое из неравенств (5.2.2).

Левое же получится аналогичным преобразованием второго неравенства в (5.2.3).

Доказательство (5.2.2) закончено.

5.3. Предельные теоремы в схеме испытаний Бернулли

Неудобство и даже невозможность применения формулы Бернулли обнаруживается в связи с необходимостью в некоторых случаях сделать предельный переход в этой формуле при $n \rightarrow \infty$. Такая необходимость возникла в связи с исследованиями в области статистики народонаселения (изучались проблемы рождаемости, смертности, подсчеты для страхования и вычисления пожизненных рент и др.), которые явились основными приложениями теории вероятностей в XVIII веке. Но и в наше время, несмотря на вычислительные возможности современной техники, для решения большинства практических задач и в теоретических исследованиях удобнее пользоваться так называемыми *предельными теоремами* схемы испытаний Бернулли. Далее

приводятся наиболее известные из них, называемые *теоремой Пуассона*, *локальной теоремой Муавра-Лапласа* и *интегральной теоремой Муавра-Лапласа* ¹³.

Теорема Пуассона

На практике часто представляет интерес случай большого числа испытаний ($n \rightarrow \infty$) и малой вероятности успеха в одном испытании ($p \rightarrow 0$) ¹⁴. Однако формула Бернулли непригодна из-за громоздких вычислений и возникающих погрешностей округления. В этом случае удобно пользоваться приближением Пуассона, которое содержится в следующей теореме.

Теорема 5.3.1. Пусть $p = p(n)$, и

$$\lim_{n \rightarrow \infty} np(n) = \lambda = \text{const} > 0.$$

Тогда

$$\lim_{n \rightarrow \infty} C_n^m p^m (1-p)^{n-m} = \frac{\lambda^m}{m!} e^{-\lambda}. \quad (5.3.1)$$

Доказательство. Для любого фиксированного n (и стало быть фиксировано число $p = p(n)$), имеем:

$$\begin{aligned} C_n^m p^m (1-p)^{n-m} &= \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \left(\cdot \frac{n^m}{n^m} \right) = \\ &= \frac{(np)^m}{m!} \frac{n(n-1) \dots (n-m+1)}{n^m} (1-p)^{n-m} = \end{aligned}$$

¹³В 1733 году в дополнении к сочинению "Учение о случаях" А. Муавр доказал локальную и интегральную предельные теоремы, при выводе которых широко использовал разложение функции в степенные ряды и формулу Стирлинга, точнее будет, если назвать ее формулой Стирлинга-Муавра. И с точки зрения самого Муавра, и с современной точки зрения, предельные теоремы Муавра-Лапласа являются непосредственным развитием закона больших чисел Я. Бернулли. Указанные теоремы были впоследствии доказаны Д. Бернулли при решении классического вопроса о соотношении рождаемости девочек и мальчиков. Еще позже теоремы были доказаны Лапласом, причем ни Д. Бернулли, ни Лаплас не ссылались на своего предшественника. По этой причине заслуга Муавра в теории вероятностей была не известна до конца XIX века.

¹⁴Вероятность редких явлений в схеме Бернулли.

$$= \frac{(np)^m}{m!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) (1-p)^{-m} \left[(1-p)^{\frac{1}{p}}\right]^{np}.$$

Заменим в полученном выражении p на $p(n)$ и перейдем к пределу при $n \rightarrow \infty$. Из условия теоремы $np(n) \rightarrow \lambda$ вытекает, что $p(n) \rightarrow 0$. Как известно

$$\lim_{\alpha \rightarrow 0} (1 - \alpha)^{1/\alpha} = e^{-1}$$

(второй замечательный предел), поэтому

$$\lim_{n \rightarrow \infty} \left[(1 - p(n))^{\frac{1}{p(n)}}\right]^{np} = e^{-\lambda}.$$

Таким образом, получаем (5.3.1).

Обычно используется теорема Пуассона, когда n велико, а p мало, в следующем виде.

$$P_n(m) \approx \frac{(np)^m}{m!} e^{-np}. \quad (5.3.1')$$

Эта формула часто называется пуассоновским приближением. Если мало число q , то пуассоновским приближением можно воспользоваться для определения вероятности числа неудач: вероятность того, что в n испытаниях событие A не произойдет $n - m$ раз, приближенно вычисляется по формуле

$$P_n(n - m) \approx \frac{(nq)^{n-m}}{(n - m)!} e^{-nq}. \quad (5.3.1'')$$

Отметим, что на практике применение формулы Пуассона (5.3.1') дает хорошее приближение формулы Бернулли (5.2.1) при $n > 50$, $p \ll 1$.

Но если параметры p и q заметно отличны от 0, т.е. близки к 1/2, то используются локальная или интегральная теоремы Лапласа.

Локальная теорема Муавра-Лапласа

Для формулировки теорем Муавра-Лапласа (5.3.2) и (5.3.3) нам понадобится понятие **равномерной сходимости** последовательности. При доказательстве же используется понятие **асимптотического равенства** функций, которое позволяет упростить и сократить само доказательство. Вначале приведем соответствующие определения (см. [10], [12]).

Последовательность $\{f_n(m)\}$, каждый член которой зависит от числа m из некоторого множества M , сходится к пределу $f(m)$ *равномерно по m на множестве M* , если для любого сколь угодно малого числа $\varepsilon > 0$ найдется номер N , такой, что для всех $n > N$ выполняется неравенство $|f_n(m) - f(m)| < \varepsilon$ сразу для всех $m \in M$.

Две функции f и g называются *асимптотически равными* (или *эквивалентными*) при $x \rightarrow x_0$, если $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 1$. Обозначение асимптотического равенства

$$f(x) \sim g(x), \quad x \rightarrow x_0$$

называется *асимптотической формулой*.

Простейшим примером асимптотического равенства служит "первый замечательный предел": $\sin x \sim x, \quad x \rightarrow 0$. Асимптотические равенства легко вытекают из представления функций формулами Тейлора: если в окрестности точки x_0

$$f(x) = P_n(x) + O(x - x_0)^{n+1},$$

где P_n — соответствующий многочлен порядка n , то

$$f(x) \sim P_n(x), \quad x \rightarrow x_0.$$

В частности $\ln(1+x) \sim x - \frac{1}{2}x^2, \quad x \rightarrow 0$.

В процессе доказательства мы также пользуемся теоремой о замене функции f , эквивалентной f_1 , при переходе к пределу (см. [12], стр. 118 или [10], стр. 153): *если $f \sim f_1, \quad x \rightarrow x_0$, то*

$$\lim_{x \rightarrow x_0} \Lambda(x) f(x) = \lim_{x \rightarrow x_0} \Lambda(x) f_1(x)$$

для любой функции Λ , определенной в окрестности точки x_0 .

Теорема 5.3.2. Пусть производится n независимых испытаний, в каждом из которых вероятность наступления некоторого события A постоянна и равна p ($0 < p < 1$), $q = 1 - p$ и m — неотрицательное число, не превосходящее n , такое, что дробь $x_m = \frac{m-np}{\sqrt{npq}}$ принадлежит отрезку $[a, b]$. Обозначим

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (5.3.2)$$

Тогда

$$P_n(m) = \frac{\varphi(x_m)}{\sqrt{npq}} (1 + \varepsilon_n), \quad \lim_{n \rightarrow \infty} \varepsilon_n = 0, \quad (5.3.3)$$

где $\varepsilon_n \rightarrow 0$ равномерно по всем m , при которых x_m принадлежит фиксированному (конечному) отрезку $[a, b]$ ¹⁵.

Д о к а з а т е л ь с т в о. Формула Стирлинга

$$n! = \sqrt{2\pi n} n^n e^{-n} e^{\Theta(n)}, \quad |\Theta(n)| \leq \frac{1}{12n}$$

дает возможность заменить все факториалы в формуле Бернулли по асимптотической формуле¹⁶

$$n! \sim \sqrt{2\pi n} n^n e^{-n} : \lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi n} n^n e^{-n}} = \lim_{n \rightarrow \infty} e^{\Theta(n)} = 1.$$

Это приводит к выражению

$$\begin{aligned} P_n(m) &= \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \sim \\ &\sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} \frac{n^n}{m^m (n-m)^{n-m}} p^m (1-p)^{n-m} = \end{aligned}$$

¹⁵Следует отметить, что в книге [8] дана наиболее ясная формулировка этой теоремы.

¹⁶Натуральные числа имеют единственную точку сгущения $+\infty$, что позволяет не дописывать в асимптотических формулах для последовательностей " $n \rightarrow +\infty$ ".

$$= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} \left(\frac{np}{m}\right)^m \left(\frac{nq}{n-m}\right)^{n-m}. \quad (5.3.4)$$

Введем обозначение

$$z = \left(\frac{np}{m}\right)^m \left(\frac{nq}{n-m}\right)^{n-m}.$$

Тогда

$$\ln z = -m \ln \left(\frac{m}{np}\right) - (n-m) \ln \frac{n-m}{nq}.$$

Из равенства $x_m = \frac{m-np}{\sqrt{npq}}$ находим

$$m = x_m \sqrt{npq} + np, \Rightarrow$$

$$n-m = n - x_m \sqrt{npq} - np = n(1-p) - x_m \sqrt{npq}. \quad (5.3.5)$$

Откуда

$$\frac{m}{np} = 1 + x_m \sqrt{\frac{q}{np}}, \quad \frac{n-m}{nq} = 1 - x_m \sqrt{\frac{p}{nq}}.$$

Поэтому

$$\begin{aligned} \ln z = & -(x_m \sqrt{npq} + np) \ln \left(1 + x_m \sqrt{\frac{q}{np}}\right) - \\ & -(n(1-p) - x_m \sqrt{npq}) \ln \left(1 - x_m \sqrt{\frac{p}{nq}}\right). \end{aligned}$$

Из формулы Маклорена для логарифмической функции вытекает асимптотическая формула

$$\ln(1+x) \sim x - \frac{1}{2}x^2, \quad x \rightarrow 0.$$

Следовательно, имеем следующее асимптотическое равенство при $n \rightarrow \infty$

$$\ln z \sim -(x_m \sqrt{npq} + np) \left(x_m \sqrt{\frac{q}{np}} - \frac{x_m^2}{2} \frac{q}{np}\right) -$$

$$\begin{aligned}
& -(nq - x_m \sqrt{npq}) \left(-x_m \sqrt{\frac{p}{nq}} - \frac{x_m^2}{2} \frac{p}{nq} \right) = \\
& = - \left[x_m^2 q + x_m^2 p - \frac{x_m^2}{2} q - \frac{x_m^2}{2} p - \frac{x_m^3}{2} \left(\frac{q^{3/2}}{\sqrt{np}} - \frac{p^{3/2}}{\sqrt{nq}} \right) \right] = \\
& = - \left[\frac{x_m^2}{2} - \frac{x_m^3}{2} \frac{q^2 - p^2}{\sqrt{npq}} \right] = -\frac{x_m^2}{2} + O\left(\frac{1}{\sqrt{n}}\right) \sim -\frac{x_m^2}{2}.
\end{aligned}$$

Таким образом,

$$\ln z \sim -\frac{x_m^2}{2} \implies z \sim e^{-\frac{x_m^2}{2}}.$$

Возвращаясь к (5.3.4), имеем:

$$P_n(m) \sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} e^{-x_m^2/2}. \quad (5.3.6)$$

Но из (5.3.5)

$$\sqrt{\frac{n}{m(n-m)}} = \sqrt{\frac{n}{(x_m \sqrt{npq} + np)(nq - x_m \sqrt{npq})}}.$$

При этом, исходя из условия $x_m \in [a; b]$, где a, b – фиксированные числа, получим:

$$\lim_{n \rightarrow \infty} \frac{x_m \sqrt{npq} + np}{np} = 1, \quad \lim_{n \rightarrow \infty} \frac{nq - x_m \sqrt{npq}}{nq} = 1,$$

т.е.

$$(x_m \sqrt{npq} + np) \sim np, \quad (nq - x_m \sqrt{npq}) \sim nq.$$

Таким образом,

$$\sqrt{\frac{n}{m(n-m)}} = \sqrt{\frac{n}{(x_m \sqrt{npq} + np)(nq - x_m \sqrt{npq})}} \sim \sqrt{\frac{n}{np \cdot nq}} = \frac{1}{\sqrt{npq}}.$$

Подставляя это значение в (5.3.6), получим утверждение теоремы.

Доказательство закончено.

Замечание 5.3.1. В доказательстве этой теоремы использовалось лишь понятие эквивалентных функций для факториалов и логарифмов, поэтому в заключение можем лишь утверждать, что $\varepsilon_n \rightarrow 0$. Но и для факториалов, и для логарифмов можно было использовать "точные" асимптотические равенства

$$n! = \sqrt{2\pi n} n^n e^{-n+O(1/12n)}, \quad n \rightarrow \infty;$$

$$\ln(1+x) = x - \frac{x^2}{2} + O\left(\frac{x^3}{3}\right), \quad x \rightarrow 0,$$

и тогда для остатка формулы (5.3.3) вытекает оценка (см. [19], [15])

$$|\varepsilon_n| = O\left(\frac{1}{\sqrt{n}}\right).$$

Замечание 5.3.2. Функция (5.3.2) табулирована; краткая таблица значений этой функции приведена в конце учебного пособия.

Интегральная теорема Муавра-Лапласа

Теорема 5.3.3 Вероятность того, что в n испытаниях Бернулли событие A ($P(A)=p$, $P(\bar{A})=q$) наступит не менее m_1 раз и не более m_2 , приближенно равна

$$P_n(m_1; m_2) = \Phi(x_{m_2}) - \Phi(x_{m_1}),$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

— функция Лапласа (называется также "функция распределения нормального закона"), $x_m = \frac{m-np}{\sqrt{npq}}$.

Доказательство. Снова рассмотрим приближенную формулу (5.3.3) предыдущей теоремы:

$$P_n(m) \approx \frac{\varphi(x_m)}{\sqrt{npq}}, \quad x_m = \frac{m-np}{\sqrt{npq}}.$$

Обратим внимание, что

$$x_{m+1} - x_m = \frac{1}{\sqrt{npq}} \rightarrow 0, \quad n \rightarrow \infty.$$

Но тогда $P_n(m_1; m_2) = P_n(m_1 \text{ или } m_1 + 1 \text{ или } \dots \text{ или } m_2) =$

$$= P_n(m_1) + P_n(m_1 + 1) + \dots + P_n(m_2) =$$

$$= \varphi(x_{m_1})\Delta x + \varphi(x_{m_1+1})\Delta x + \dots + \varphi(x_{m_2})\Delta x, \quad \Delta x = \frac{1}{\sqrt{npq}}.$$

Последнее выражение представляет собой одну из интегральных сумм Римана для интеграла

$$\int_{\frac{m_1 - np}{\sqrt{npq}}}^{\frac{m_2 - np}{\sqrt{npq}}} \varphi(x) dx.$$

Утверждение теоремы вытекает теперь из того, что функция $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ интегрируема (в несобственном смысле) по R_1 , и следовательно, по любому интервалу, принадлежащему R_1 .

Замечание 5.3.3. Решение многих практических задач требует умения вычислять значение интеграла $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ при любых значениях x . Для расчетов используют специальные таблицы значений функции $\Phi(x)$. Краткая таблица значений этой функции приведена в конце книги. Причем таблица функции $\Phi(x)$ составлена для положительных x ; для отрицательных x функция $\Phi(x)$ определяется из равенства $\Phi(-x) = -\Phi(x)$.

Глава 2

Случайные величины. Функции распределения

Следующим фундаментальным понятием теории вероятностей является понятие случайной величины. Во многих практических задачах роль "случайного события" играют числа (время, деньги, килограммы и т.д.). Работа с числами позволяет увидеть в "бытовом тексте" математическую задачу в упрощенной форме. Это, разумеется, привлекло внимание исследователей. Но есть и другое принципиальное значение для теоретических исследований — среди чисел можно отыскать среднюю величину, что невозможно среди "не числовых событий". Например, какое среднее событие между "орел" и "решка"? Но если "орел" - нуль копеек, "решка" - одна копейка выигрыша, то среднее 0,5 - это число, откровенно указывающее игрокам в "орлянку", что выигрыш и проигрыш есть величина близкая к копейке. Далее мы увидим, как переход от "случайного события" к "случайной величине" несет много полезной информации об изучаемом процессе или явлении.

На разных этапах развития теории вероятностей понятие случайной величины играло свою роль в формировании научного

знания. В период зарождения теории вероятностей этот термин не был введен в научный обиход, но уже использовался многими учеными при решении практических задач. Это были задачи, возникшие в статистических исследованиях при изучении демографических проблем, для обработки астрономических наблюдений, а также задачи из области различных азартных игр, которые явились мощным стимулом развития теории вероятностей. На данном этапе следует отметить работы Д. Граунта (1620–1675) и В. Петти (1623–1687), Э. Галлея (1656–1742) и др.. Большой вклад в развитие понятия случайной величины внес Х. Гюйгенс (1629–1695)— автор первого практического руководства по теории вероятностей. В книге "О расчетах в азартной игре" (1657) рассматриваются различные примеры случайных величин: количество недостающих партий или разное число игроков в задаче о разделении ставок, возможные выигрыши в других азартных играх и т. д.. В начале книги Х. Гюйгенс вводит определение математического ожидания, которое явилось первым теоретико-вероятностным понятием. Анализируя эмпирические данные, полученные Граунтом (использовались акты регистрации смертности в Лондоне с XVI века), Гюйгенс строит кривую плотности распределения, где вероятностями служили статистические частоты.

Дальнейший этап формирования понятия случайной величины — это "искусство умозаключений" Я. Бернулли (1654–1705). Я. Бернулли рассмотрел случайную величину — число появлений интересующего события в n независимых испытаниях, которая может принимать значения $0, 1, 2, \dots, n$ с определенными вероятностями. Сегодня это пример случайной величины, распределенной по биномиальному закону.

Его племянник Н. Бернулли (1687–1753) вывел формулу математического ожидания длины случайного интервала AB , образованного фиксированной точкой A и самой правой из случайных точек B_i , $i = 1, 2, \dots, n$, равномерно распределенных на заданном отрезке AB . Эта задача была первой, в которой было введено равномерное распределение. В классической задаче о соотношении мальчиков и девочек Н. Бернулли определил ве-

роятность нахождения в заданных пределах числа ежегодного рождения мальчиков. Необходимо отметить, что данный период имеет огромное значение для дальнейшего развития всей теории вероятностей.

Я. Бернулли и его последователями были получены значимые результаты о представлении и числовых характеристиках случайной величины, но основной заслугой явилась доказанная им теорема, которая носит его имя. Она явилась первой в цепи предположений, образующих закон больших чисел (следует заметить, что термин "закон больших чисел" был впервые введен С. Пуассоном в 1835 г.).

Третий этап — этап развития математической теории ошибок наблюдений — связан с такими именами, как Р. Котс (1682–1716), Т. Симсон (1710–1761), Д. Бернулли (1700–1782), П. Лаплас (1749–1827). Исследуя свойства ошибок, оценку точности измерений, закономерности распределения ошибок, ученые используют понятие плотности распределения и различные виды законов распределения, но нигде не вводят понятия случайной величины.

Завершающий этап — период непосредственного формирования понятия случайной величины и ее характеристик. Первый шаг в этом направлении был сделан С. Пуассоном (1781–1840) в работе "О вероятности средних результатов наблюдений" (1832), в которой он пишет о некоторой вещи, способной принимать значения $a_1, a_2, \dots, a_\lambda$ соответственно с вероятностями $p_1, p_2, \dots, p_\lambda$. Термин "вещь" вскоре был заменен П.Л. Чебышевым на понятие "величина". В работах А.М. Ляпунова (1857–1918) систематически используется термин "случайная величина". В работе "Об одном приложении теории вероятностей" А.М. Ляпунов определил функцию распределения в том виде, в котором мы используем ее сейчас. Понятие случайной величины, данное Пуассоном, сегодня нельзя считать математическим, оно скорее носит описательный характер. Строго формализованное определение случайной величины было дано А.Н. Колмогоровым (1903–1987) в начале прошлого столетия в знаменитой книге "Основные понятия теории вероятностей" (1933).

Перейдем к математическому описанию понятия *случайной величины*.

§ 6. Случайные величины

Рассмотрим некоторое вероятностное пространство (Ω, \mathcal{U}, P) . Каждому элементарному исходу ω_i поставим в соответствие некоторое число x_i . Тем самым на множестве Ω определяется числовая функция, которую обозначим $X = X(\omega)$ от элементарного события $\omega \in \Omega$ и называется *случайной величиной*, если для любого числа $x \in (-\infty; +\infty)$ выполнено $\{X < x\} \in \mathcal{U}$. Пространство элементарных исходов заменяется соответствующим множеством *случайных значений* на прямой R_1 , которое удобно обозначать тем же символом Ω и называть множеством *элементарных реализаций случайной величины* X .

Если числовое множество Ω состоит из изолированных точек, т.е. $\forall x_i \in \Omega$ существует полная окрестность x_i , не содержащая других точек из Ω , то случайная величина X называется *дискретной случайной величиной*.

Число возможных значений дискретной случайной величины может быть конечным или бесконечным. Например, число очков, выпавших на верхней грани при одном бросании игральной кости равно 6; количество вызовов на телефонной станции — любое число из натурального ряда чисел.

Если же Ω состоит из сплошного промежутка (отрезка, интервала, возможно, бесконечных), то случайная величина X называется *непрерывной случайной величиной*.

Алгебра случайных событий \mathcal{U} , таким образом, заменяется алгеброй случайных значений, для которой будем использовать тот же символ \mathcal{U} .

Случайные величины будем обозначать большими буквами латинского алфавита X, Y, Z, \dots , а их возможные значения (ре-

ализации) — соответствующими малыми буквами x, y, z, \dots .

Законом распределения случайной величины X будем называть вероятность $P\{X \in B\}$, рассматриваемую как функцию числового множества B , если $\{X \in B\} \in \mathcal{U}$.

В этом определении множество B произвольное (числовое) подмножество Ω , и поэтому само определение употребляется лишь в случаях, когда оно фиксировано.

Пусть дискретная случайная величина задана с помощью таблицы, верхний ряд которой состоит из чисел x_i , а числа нижнего ряда — их вероятности:

X	x_1	x_2	x_3	\dots	x_n
P	p_1	p_2	p_3	\dots	p_n

где

$$p_i \geq 0, \quad \sum_{i=1}^n p_i = 1.$$

Эта таблица называется законом распределения дискретной случайной величины. Если $B = \{x_{i_k}\}_{k=\overline{1,m}}$, то

$$P(X \in B) = \sum_{k=1}^m P(x_{i_k}).$$

Аналогичное понятие для непрерывных случайных величин удобно связать с множеством $B = (-\infty, x)$, и этот закон называется "функцией распределения вероятностей случайной величины".

6.1. Функция распределения вероятности

Вначале выясним, что собственно должно обозначать слово "распределение" и каким образом этому физическому термину можно придать математический смысл. Представим длинную проволоку (материальную нить). "Распределение массы по длине" естественно задать с помощью функции $M = M(l)$, аргумент которой l — длина соответствующего куска проволоки.

Если масса распределена равномерно, то $M(l) = k \cdot l$, то есть любые одинакового размера l куски имеют равные массы. Но приведенная модель "распределения" годится лишь в случае, когда одинаковые куски проволоки весят одинаково, т.е. масса по проволоке "распределена равномерно". Теперь представим, что одинаковые куски имеют неодинаковые массы. Тогда для описания "распределения массы" потребуется фиксация самих кусков. Это можно сделать множеством способов, например, раскрашивая куски в разные цвета, а самый простой — введения координат на проволоке, и тогда любой кусок получает имя — кусок $[a, b]$. В этом случае распределение массы удобно задать функцией $M = M(l < x)$ — масса участка длиной x , измеряемого от конца, помещенного в начало координат¹. Это в самом деле удобно, поскольку масса любого куска "по имени $[a, b]$ ", т.е. с координатами концов $x = a$ и $x = b$, определяется по очевидной и очень простой формуле $M(a < l < b) = M(l < b) - M(l < a)$.

Этими рассуждениями естественно воспользоваться как способом задания вероятностной характеристики случайной величины (с той лишь разницей, что случайные величины могут принимать произвольные значения, в том числе и отрицательные, и их удобнее считать принадлежащей всей прямой $(-\infty, +\infty)$, а не полупрямой $[0, +\infty)$, как в описании распределения массы по длине проволоки).

Определение 6.1. *Функцией распределения вероятностей случайной величины X называется вероятность того, что случайная величина X примет значение из интервала $[-\infty, x)$:*

$$F(x) = P\{X < x\}.$$

Простейший пример.

♦ 1. Подбрасывается один раз монета. Пространство элементарных исходов $\Omega = \{O, R\}$, алгебра \mathcal{U} — минимальная, состоя-

¹Следует писать $M = M(x)$, но в точке $l = x$ определены две массы — слева $l < x$ и справа $l > x$. Неравенство нам потребовалось для фиксации одного из кусков проволоки и для согласования с последующим текстом, где "распределение" определяется именно как функция от неравенства: вероятность попадания в точку слева от точки x .

щая из Ω и пустого множества \emptyset , куда входят все события типа "не упала", "упала и пропала", "встала на ребро", "упала сразу двумя сторонами вверх" и т.д.. Вероятность на элементах алгебры определяется как на равновозможных исходах, т.е. $P(O) = P(R) = \frac{1}{2}$. Каждому элементарному исходу произвольно приписываем числовые значения², например, "O"— π , "R"—7, и тем самым получаем случайную величину $X = \{\pi, 7\}$, каждое значение которой принимается с вероятностью $\frac{1}{2}$. Ясно, что событие $\{X < \pi\}$ — пустое множество с вероятностью $P\{X < x, -\infty < x < \pi\} = 0$, событие $\{X < 7\}$ состоит из одного значения $\{X = \pi\}$, поэтому $P\{X < x, \pi \leq x < 7\} = \frac{1}{2}$ и, наконец, событие $\{X < x = 7 + t, t \geq 0\}$ состоит из двух точек π и 7, а вероятность $P\{X < x = 7 + t\} = P\{\pi \text{ или } 7\} = 1$. Таким образом, в этом случае функция распределения $F(x)$ — ступенчатая функция (непрерывная слева в точках разрыва), график которой изображен на рисунке 8.

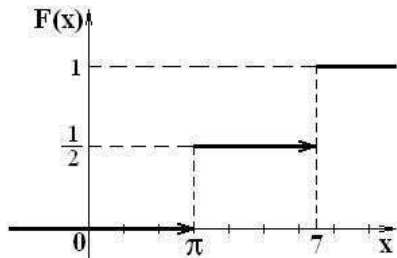


Рис. 8

♦ **2.** Случайная величина принимает значения x на отрезке $[a, a + 1]$ с вероятностью, равной расстоянию от точки a до точки x . В этом случае функцией распределения является непрерывная функция

$$F(x) = \begin{cases} 0, & x < a, \\ x - a, & x \in [a, a + 1], \\ 1, & x \geq a + 1, \end{cases}$$

²Чаще в этом случае приписываются значения -1 и +1, символизирующие проигрыш и выигрыш соответственно. Но по теории выбирать можно любые числовые значения, сколь нелепыми они не покажутся.

график которой изображен на рисунке 9.

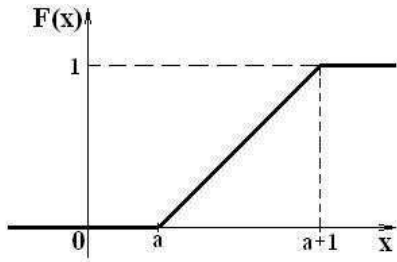


Рис. 9

Свойства функции распределения F .

1. $0 \leq F(x) \leq 1$.

2. $F(x)$ — неубывающая функция, непрерывная слева, то есть

$$F(x_2) \geq F(x_1), \quad x_2 > x_1; \quad \lim_{x \rightarrow x_0 - 0} F(x) = F(x_0).$$

3. $\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1$.

Если же возможные значения случайной величины X принадлежат интервалу $(a; b)$, то $F(x) = 0$, если $x \leq a$ и $F(x) = 1$, если $x \geq b$.

Доказательство этих свойств с очевидностью следует из определения функции распределения $F(x)$.

Функция распределения используется при решении главной задачи теории случайных величин: найти вероятность, что случайная величина попала в заданный интервал. Справедлива формула

$$P\{x_1 \leq X < x_2\} = F(x_2) - F(x_1),$$

а если функция $F(x)$ непрерывна, то

$$P\{x_1 < X < x_2\} = F(x_2) - F(x_1); \quad (6.1.1)$$

Доказательство формулы (6.1.1). Заметим, что $\{X < x_2\} = \{x_1 \leq X < x_2\} \cup \{X < x_1\}$. В равенстве справа — объединение непересекающихся множеств, которое представляет

сумму несовместных событий. По аксиоме АЗ, получим $P\{X \leq x_2\} = P\{x_1 \leq X < x_2\} + P\{X < x_1\}$. Откуда

$$P\{x_1 \leq X < x_2\} = P\{X < x_2\} - P\{X < x_1\} = F(x_2) - F(x_1). \quad (6.1.2)$$

Если же функция $F(x)$ непрерывна в окрестности точки x_1 , то вероятность принять значение $x = x_1$ находится предельным переходом

$$\begin{aligned} P\{X = x_1\} &= \lim_{\Delta x \rightarrow 0} P\{x_1 \leq X < x_1 + \Delta x\} = \\ &= \lim_{\Delta x \rightarrow 0} (F(x_1 + \Delta x) - F(x_1)) = 0, \end{aligned}$$

т.е., если функция распределения вероятности F непрерывна, то вероятность принять значение в фиксированной точке равна нулю³. В этом случае слева в равенствах (6.1.2) можно записать строгое неравенство, и мы получим (6.1.1).

Доказательство закончено.

Дискретная случайная величина. Пусть дискретная случайная величина задана законом распределения

X	x_1	x_2	x_3	\dots	x_n
P	p_1	p_2	p_3	\dots	p_n

, $x_i < x_{i+1}$, $\sum_{i=1}^n p_i = 1$.

По определению $F(x) = P\{X < x\}$. Для дискретной случайной величины принять значение меньше, чем x , означает, что X принимает только те x_i , которые меньше x . Поэтому, если $X < x_1$, то очевидно $F(x) = 0$, так как не может случайная величина принять значение, меньшее самого наименьшего;

если $x_1 \leq X < x_2$, то такое значение всего одно, а именно x_1 , следовательно, $F(x) = p_1$;

если $x_2 \leq X < x_3$, то таких значений два: x_1 и x_2 , следовательно, $F(x) = P\{X = x_1 \text{ или } X = x_2\} = P\{X = x_1\} + P\{X = x_2\} = p_1 + p_2$ и так далее.

³В этом заключается "парадокс непрерывности": вероятность равна нулю, значит, событие $\{X = x_0\}$ невозможное, но какое-то значение в результате испытания случайная величина X обязательно примет, поэтому именно оно не невозможное!?

В результате мы получаем ступенчатую функцию

$$F(x) = \sum_{i: x_i < x} p_i, \quad (6.1.2)$$

непрерывную слева, изображенную на рисунке 10.

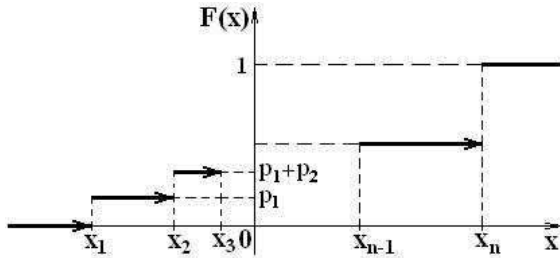


Рис. 10

Удобно эту функцию распределения записывать с помощью функции Хевисайда (функция включения):

$$\Theta(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 & , \quad x \geq 0 \end{cases}$$

следующим образом

$$F(x) = \sum_{i=1}^n p_i \Theta(x - x_i). \quad (6.1.3)$$

6.2. Плотность распределения вероятности

"Плотность" как физическая величина хорошо известна и не требует пояснений. "Плотность вероятности" определяется по полной аналогии с "физической". Начнем наши объяснения с наиболее простого и классического — плотности вероятностей случайной величины, распределенной непрерывно.

Непрерывная случайная величина. Величину

$$\frac{P\{x < X < x + \Delta x\}}{\Delta x}$$

естественно назвать средней плотностью распределения вероятностей случайной величины X на интервале $(x, x + \Delta x)$.

Как нам известно (формула (6.1.1)),

$$P\{x < X < x + \Delta x\} = F(x + \Delta x) - F(x) = \Delta F,$$

поэтому, если функция $F(x)$ в точке x имеет производную, то определена функция

$$p(x) = \lim_{\Delta x \rightarrow 0} \frac{P\{x < X < x + \Delta x\}}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta F}{\Delta x} = F'(x), \quad (6.2.1)$$

которая и представляет собой *плотность распределения вероятности случайной величины X* с дифференцируемой в точке x функцией распределения F . Отметим, что если $F(x)$ имеет производную в любой точке, то равенство (6.2.1) говорит о том, что функция $p(x)$ является первообразной к $F(x)$, следовательно,

$$F(x) = \int_{-\infty}^x p(t) dt, \quad (6.2.2)$$

и имеет место формула Ньютона-Лейбница

$$P\{a < X < b\} = \int_a^b p(x) dx = F(b) - F(a). \quad (6.2.3)$$

Из (6.2.1) вытекает также, что функция p непрерывна, поэтому из теоремы о представлении функции в окрестности предельной точки следует асимптотическое равенство (эквивалентность)

$$P\{x < X < x + \Delta x\} \sim p(x) \Delta x, \quad \Delta x \rightarrow 0.$$

Впрочем, последнее вытекает также из (6.2.2) по свойству "среднего значения" определенного интеграла.

В рамках непрерывно дифференцируемых функций равенства (6.2.1) и (6.2.2) равносильны, и любое из них можно принять за определение плотности распределения вероятностей случайной величины.

Но в (6.2.1) плотность $p(x)$ определена лишь при условии, что распределение вероятностей представлено дифференцируемой функцией $F(x)$, а в (6.2.2) функция распределения определена при условии, что функция $p(x)$ всего лишь интегрируема по Лебегу (или по Риману, но тогда в несобственном смысле). Во втором случае можно утверждать, что функция $F(x)$ непрерывна⁴, и нельзя утверждать, что она дифференцируема в каждой точке соответствующего интервала, как это требовалось в первом случае. Из сказанного вытекает, что равенство (6.2.2) более общее, следовательно, именно его следует взять за определение *плотности распределения*, но при двух дополнительных условиях:

функция $p = p(x)$ неотрицательна, поскольку должно выполняться главное свойство плотности — $p(x) \Delta x \sim P\{X \in [x, x + \Delta x]\}$, $\Delta x \rightarrow 0$, а $P\{X \in [x, x + \Delta x]\} \geq 0$;
выполняется равенство

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

означающее, что событие (см. (6.2.3)) $\{X \in [-\infty, +\infty]\}$ — достоверное.

Отметим, что функции вида (6.2.2) называются *абсолютно непрерывными*⁵. Они обладают очень важным для нас свой-

⁴Непрерывность интеграла как функции верхнего предела для R-интегрируемых функций описана во всех учебниках по математическому анализу. Для L-интегралов это тоже так, но теория значительно богаче и приводит к понятию *абсолютной непрерывности*.

⁵Вообще говоря, функция F называется абсолютно непрерывной (на $[a, b]$), если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любой системы попарно не пересекающихся интервалов $(\in [a, b])$ суммарной длины меньше δ , сумма абсолютных величин приращений функции F на этих интервалах будет меньше, чем ε . Это равносильно тому, что найдется L-интегрируемая функция $p(x)$ такая, что $F(x) = A + \int_{-\infty}^x p(t) dt$. Сведения об абсолютно непрерывных функциях можно найти в курсах математического анализа и теории функций действительного переменного, см. например [12],

ством: равенство

$$F'(x) = \left(\int_{-\infty}^x p(t) dt \right)' = p(x)$$

выполняется почти всюду и всегда в смысле "обобщенных функций"⁶.

Введем следующее определение.

Определение 6.2.1 Случайная величина X называется *распределенной абсолютно непрерывно*, если существует неотрицательная L -интегрируемая функция $p(x)$, такая, что при любом x функция распределения вероятностей F представляется в виде

$$F(x) = \int_{-\infty}^x p(x) dx, \quad \int_{-\infty}^{+\infty} p(x) dx = 1 \quad (6.2.4)$$

При этом функция $p(x)$ называется *плотностью распределения вероятностей*⁷.

Еще раз подчеркнем, что в определении (6.2.4) функция $p(x)$ не обязана быть определенной в каждой точке (согласно теореме Лебега, см. например [12], т.1, стр.343), интегрируемая функция может быть не определена на множестве лебеговой меры нуль. Функция же F , напротив, определена и непрерывна в каждой точке. Но главное, что делает определение удобным, это равенство

$$F'(x) = p(x)$$

т.2, стр. 373-375, непосредственно для неопределенных интегралов Лебега соответствующее доказательство содержится в книге [22], стр. 169-170.

⁶Производная обобщенной функции, или производная распределения, определяется далее в постулате ii. Равенство $F' = p$ показывает, что *обобщенная производная от абсолютно непрерывных функций эквивалентна обычной*. А это уже означает: определение (6.2.1) эквивалентно равенству (6.2.1).

⁷ L -интегрируемая функция — функция, интегрируемая по Лебегу, см. [12], т. 2, гл. 19. В случае, если читатель не знаком или плохо знаком с интегралом Лебега, рекомендуем считать, что в равенстве (6.2.4) интеграл Римана, понимаемый как несобственный интеграл.

выполняется почти всюду и делает определения (6.2.1) и (6.2.4) "почти" равносильными.

В этой связи приведем пример.

♦ **Пример 1.** Найти константу A , при которой функция

$$p(x) = \begin{cases} 0, & x \leq 0, \\ \frac{A}{\sqrt{x}}, & x \in (0, 1], \\ 0, & x > 1. \end{cases}$$

является плотностью распределения некоторой случайной величины X . Найти ее функцию распределения $F(x)$ и вероятность того, что случайная величина X примет значения из интервала $(-\frac{1}{2}; \frac{1}{2})$.

Решение. Используя свойство (см. (6.2.4)) $\int_{-\infty}^{+\infty} p(x) dx = 1$, мы получим $A = \frac{1}{2}$. Следовательно,

$$F(x) = \int_{-\infty}^x p(x) dx = \begin{cases} 0, & x \leq 0, \\ \sqrt{x}, & x \in (0, 1], \\ 1, & x > 1 \end{cases}$$

Как видим, эта функция непрерывна, но не дифференцируема в двух точках $x = 0$ и $x = 1$. Тем не менее мы можем решить основную задачу теории случайных величин — определить вероятность попадания случайной величины в заданный интервал. Например, вероятность попасть в интервал $(-\frac{1}{2}, \frac{1}{2})$ равна

$$P\left\{-\frac{1}{2} < X < \frac{1}{2}\right\} = F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right) = F\left(\frac{1}{2}\right) = \frac{1}{\sqrt{2}}.$$

Замечание 6.2.1. Геометрически свойство $\int_{-\infty}^{+\infty} p(x) dx = 1$ означает, что вся площадь криволинейной трапеции, ограниченной осью OX и кривой плотности распределения $p(x)$, равна 1.

Замечание 6.2.2. Если все возможные значения случайной величины X принадлежат $(a; b)$, то $\int_{-\infty}^{+\infty} p(x) dx = \int_a^b p(x) dx = 1$.

Дискретная случайная величина

Для дискретной случайной величины, казалось бы, невозможно ввести понятие плотности распределения вероятности аналогичное (6.2.2) как для непрерывных случайных величин. Это так, если на плотность $p(x)$ смотреть как на функцию. Но можно воспользоваться теорией обобщенных функций⁸ для определения функции распределения вероятности F в виде "сингулярного функционала".

Соответствующий математический аппарат (введен П.Дираком в 1925 г., а его корректное математическое описание дано С.Л.Соболевым в 30-х и Л.Шварцем в 50-х годах XX столетия) приспособлен для интегрального описания масс, зарядов, плотностей (масс) и др., сосредоточенных в точке (см. Дополнение 1).

Н е б о л ь ш о е в с т у п л е н и е. Каждый, изучающий физику и математику, наверное, хотя бы один раз задал себе вопрос о сущности математического предельного перехода в физике. Например, что такое скорость в момент времени: времени в пути — нет, пройденного расстояния — нет, а скорость есть. О плотности в точке, разумеется, то же самое — материи нет, объема, где эта материя обитает, тоже нет, а плотность есть. Ответ прост. Все дело в "среднем значении". Например, если в 1 см^3 находится какая-то масса вещества, то, поделив объем на 10, получим и в 10 раз меньше вещества, поделив на сто, — в сто раз меньше вещества и т.д., неизменным остается именно физическая величина, равная отношению "массы в объеме к этому объему", — плотность. Если же распределение неравномерное, то вычисляется множество "средних значений", и предел средних — опять же константа, характеризующая физическую инварианту в "очень маленьком объеме", которую по аналогии с равномерным распределением на-

⁸Теорию линейных непрерывных функционалов (распределений, обобщенных функций) содержат все современные учебники по математической физике (см. например [3]), поскольку в рамках этой теории удастся ставить и решать физические и инженерные задачи в наиболее натуральной форме. Для наших целей необходима очень малая часть этой теории, которая заключается в определении δ -распределения и двух постулатах. Исторические справки, пояснения к определению δ -распределения содержит Дополнение 1, которое рекомендуем прочитать не знакомым с этой теорией.

зывают плотностью. Мы тоже исходим из понятия среднего, используя введенное П. Дираком δ -распределение, но поступая, как математики, заменив "среднее арифметическое" — "среднеинтегральным", считая, что плотности, как функции точки нет, но есть распределение, определяемое через интегральное среднее функций, образующих δ -образную последовательность.

Приведем эти рассуждения. Пусть дискретная случайная величина X задана законом распределения $P\{X \in B\}$. И пусть значению случайной величины x_i отвечает вероятность p_i . Размажем вероятность вокруг точки $x = x_i$, например, равномерно в ε -окрестности точки x_i с помощью функции

$$\delta_\varepsilon(x - x_i) = \begin{cases} \frac{1}{2\varepsilon}, & -\varepsilon \leq x - x_i \leq \varepsilon \\ 0, & |x - x_i| > \varepsilon \end{cases}$$

или неравномерно, но с носителем в ε -окрестности точки x_i и с условием, что $\int_{-\infty}^{+\infty} \delta_\varepsilon(x) dx = \int_{-\varepsilon}^{+\varepsilon} \delta_\varepsilon(x) dx = 1$.

Положим $p_\varepsilon(x) = \sum_{i=1}^n p_i \delta_\varepsilon(x - x_i)$. Тогда, очевидно

$$F(x) = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^x p_\varepsilon(t) dt = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^x \sum_{i=1}^n p_i \delta_\varepsilon(t - x_i) dt = \sum_{i: x_i < x} p_i,$$

т.е. мы получили формулу (6.1.2). Далее, следуя Дираку, необходимо спрятать предельные переходы в обозначение, положив плотность распределения, равную

$$p(x) = \sum_i p_i \delta(x - x_i),$$

а "действие" $\delta(x)$ на всякую непрерывную функцию f задать равенством $\int_{-\infty}^{+\infty} f(x) \delta(x) dx = f(0)$ или

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1, \quad \int_{-\infty}^{+\infty} \delta(t - x) f(t) dt = \int_{-\infty}^{+\infty} \delta(t) f(x + t) dt = f(x).$$

Обоснование этих равенств вытекает из первой теоремы о среднем значении (для простоты интегралы понимаем в смысле главного значения):

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} f(t) \delta_{\varepsilon}(x-t) dt &= \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} f(x+t) \delta_{\varepsilon}(t) dt = \lim_{\varepsilon \rightarrow 0} f(\xi) \int_{-\varepsilon}^{\varepsilon} \delta_{\varepsilon} dx = \\ &= \lim_{\varepsilon \rightarrow 0} f(\xi) = f(x). \end{aligned}$$

Тем самым мы получили основание использовать для дискретных случайных величин те же формулы, что и для абсолютно непрерывно распределенных случайных величин. Теория о распределении дискретных случайных величин выглядит следующим образом.

Пусть $\delta = \delta(x)$ — дельта-распределение (обобщенная δ -функция Дирака), представляющее собой функционал⁹, ставящий в соответствие каждой функции f , непрерывной в окрестности начала координат, ее значение в нуле:

$$(\delta, f) = \int_{-\infty}^{\infty} f(x) \delta(x) dx \stackrel{\text{def}}{=} f(0).$$

Тогда плотность случайной величины X , заданной законом

X	x_1	x_2	x_3	\dots	x_n
P	p_1	p_2	p_3	\dots	p_n

(6.2.5)

определяется по формуле

$$p(x) = \sum_{i=1}^n p_i \delta(x - x_i) \quad (6.2.5')$$

и называется *сингулярным распределением* (сингулярной обобщенной функцией).

⁹Функционал f , определенный на множестве функций $\{\varphi\}$, принято обозначать, как скалярное произведение (f, φ) . Множество $\{\varphi\}$ называется классом основных или пробных функций и состоит из таких функций, для которых этот функционал определен. Для нашей цели достаточно считать, что это интегрируемые, непрерывно дифференцируемые функции.

Отметим, что закон распределения, определенный таблицей (6.2.5), равносильно выражению (6.2.5'), поэтому выражение (6.2.5'), как и таблицу (6.2.5), будем также называть *законом распределения дискретной случайной величины X* .

О корректности определения плотности (закона) по формуле (6.2.5'). Действительно, функция распределения должна быть определена интегралом от плотности. Имеем:

$$F(x) = P\{X < x\} = \int_{-\infty}^x p(t) dt = \int_{-\infty}^x \sum_{i: x_i < x} p_i \delta(t - x_i) dt = \sum_{i: x_i < x} p_i,$$

или

$$F(x) = \sum_{i=1}^n p_i \Theta(x - x_i), \quad (6.2.6)$$

т.е. мы получили (6.1.3), как и в случае обычного понимания функции распределения.

Но, как мы сейчас проверим, и производная от функции распределения (в рамках теории распределений, см. далее постулат ii) (6.2.7) тоже даст плотность распределения вероятности, определенной δ -функционалом по формуле (6.2.6):

$$F'(x) = \frac{d}{dx} \sum_i p_i \Theta(x - x_i) = \sum_{i: x_i < x} p_i \delta(x - x_i). \quad (6.2.7)$$

Для доказательства (6.2.7) мы должны принять два постулата.

i) (о равенстве) если $(f, \varphi) = (g, \varphi)$ справедливо для всех функций φ , принадлежащих основному классу, то $f = g$;

ii) (о дифференцировании функционалов) $(f', \varphi) = -(f, \varphi')$, где

$$(f, \varphi) = \int_{-\infty}^{+\infty} f(x) \varphi(x) dx.$$

Заметим, что из i) следует, что функции f и g могут не совпадать только на множестве лебеговой меры нуль, а если f — дифференцируемая функция, то ii) легко проверяется интегрированием по частям.

Согласно постулата i), достаточно доказать равенство

$$(\Theta', \varphi) = (\delta, \varphi) = \varphi(0). \quad (6.2.8)$$

Имеем:

$$\begin{aligned} (\Theta', \varphi) &= -(\Theta, \varphi') = - \int_{-\infty}^{+\infty} \Theta(x) \varphi'(x) dx = \\ &= - \left[\int_{-\infty}^0 \Theta(x) \varphi'(x) dx + \int_0^{+\infty} \Theta(x) \varphi'(x) dx \right] = \\ &= - \Theta(x) \varphi(x) \Big|_{-\infty}^0 + \int_{-\infty}^0 \Theta'(x) \varphi(x) dx - \Theta(x) \varphi(x) \Big|_0^{+\infty} + \\ &\quad + \int_0^{+\infty} \Theta'(x) \varphi(x) dx. \end{aligned}$$

Здесь под знаком интеграла стоят производные от констант, поэтому интегралы исчезают. Кроме того $\varphi(\pm\infty) = 0$, и мы получим:

$$(\Theta', \varphi) = \Theta(+0) \varphi(0) - \Theta(-0) \varphi(0) = (\Theta(+0) - \Theta(-0)) \varphi(0).$$

Согласно определению функции Хевисайда, $\Theta(+0)=1$, $\Theta(-0)=0$, и тогда $(\Theta', \varphi) = \varphi(0)$. Таким образом, равенство (6.2.8) доказано, следовательно,

$$\Theta'(x) = \delta(x). \quad (6.2.9)$$

Разумеется, это равенство надо понимать в смысле принятого нами постулата i), т.е. $(\Theta'(x), \varphi) = (\delta(x), \varphi)$ для любой основной функции φ .

Теперь, используя формулу (6.2.9), получим равенство, понимаемое в смысле постулата i) и определяющее сингулярную обобщенную функцию

$$p(x) = F'(x) = \sum_{i: x_i < x}^n p_i \Theta'(x - x_i) = \sum_{i: x_i < x}^n p_i \delta(x - x_i),$$

которая называется *обобщенной плотностью распределения* дискретной случайной величины X .

С помощью обобщенной плотности распределения дискретной случайной величины удобно описывать, так называемые, дискретно-непрерывные законы распределений.

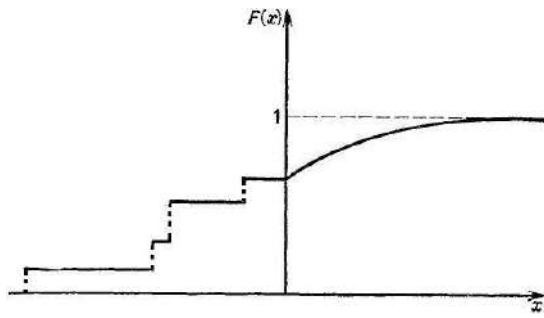


Рис. 11

Классическим примером такого распределения является распределение вероятности энергии в квантовой теории (см. [14]): после поглощения одного кванта энергии атом может перейти в одно из "возбужденных состояний", при

этом распределение вероятности энергии E частично дискретно, а частично непрерывно, как на рис. 11, где x_0 — энергия ионизации.

Функция распределения $F(x) = P\{X < x\}$ является ступенчатой при $x < x_0$ и непрерывной при $x > x_0$. Плотность распределения удобно представить в виде

$$f(x) = \sum_{i=1}^n p_i \delta(x - x_i) + f_2(x),$$

где $f_2(x) = 0$ при $x > x_0$.

§ 7. Числовые характеристики случайных величин

7.1. Математическое ожидание

Не всегда требуется полная информация о случайной величине X , выражающаяся в ее функции распределения $F(x)$. В некоторых практических задачах достаточно знать, где находится область "типичных" значений X . Одной из важнейших

характеристик "центра" этой области является математическое ожидание.

Определение 7.1.1. Математическим ожиданием случайной величины X с плотностью распределения $p = p(x)$ называется число

$$EX = \int_{-\infty}^{\infty} x p(x) dx. \quad (7.1.1)$$

Рассмотрим случай, когда X — дискретная случайная величина. Тогда плотность вероятности p оказывается законом распределения (6.2.5'): $p(x) = \sum_{i=1}^n p_i \delta(x - x_i)$, а определение (7.1.1) дает формулу

$$EX = \int_{-\infty}^{\infty} x \sum_{i=1}^n p_i \delta(x - x_i) dx = \sum_{i=1}^n x_i p_i, \quad (7.1.2)$$

определяющую математическое ожидание дискретной случайной величины X , заданной законом $p(x) = \sum_{i=1}^n p_i \delta(x - x_i)$.

Из приложения определенного интеграла к задачам механики хорошо известна формула для определения центра масс на оси x (статический момент относительно оси y) криволинейной пластины с границами $y = f(x)$ и осью x :

$$x_{c.t.} = \frac{\int_{-\infty}^{+\infty} x f(x) dx}{\int_{-\infty}^{+\infty} f(x) dx}.$$

Если в этой формуле функцию f интерпретировать как плотность распределения случайной величины X (при этом $\int_{-\infty}^{+\infty} f(x) dx = 1$), то

$$x_{c.t.} = \int_{-\infty}^{+\infty} x f(x) dx,$$

т.е. центр тяжести криволинейной пластины интерпретируется как математическое ожидание (7.1.1) случайной величины X с плотностью распределения p , равной уравнению криволинейной границы $p = f$ пластины.

Формула (7.1.2) также имеет простую механическую (и геометрическую) интерпретацию. Из аналитической геометрии известно, что система материальных точек на оси (т.е. точек с координатами x_i , в которые помещены массы m_i) имеет центр тяжести, определяемый по формуле

$$x_{c.t.} = \frac{\sum x_i m_i}{\sum m_i} = \frac{\sum x_i m_i}{m} = \sum x_i \frac{m_i}{m}.$$

Но $0 < \frac{m_i}{m} < 1$ и $\sum \frac{m_i}{m} = 1$, т.е. число $\frac{m_i}{m}$ можно интерпретировать как вероятность. Но тогда центр масс (сравните с (7.1.2)) — математическое ожидание.

В этой связи математическое ожидание случайной величины X называют *центром распределения* случайной величины.

Чтобы понять, чем отличается "математическое ожидание" от бытового "ожидания", рассмотрим пример.

Предположим, что Вы имеете n однотипных деталей и точный прибор для определения одного из размеров детали. Проведя измерения, Вы убедились, что все размеры разные. Возникает вопрос: какой размер стремился выполнить мастер при изготовлении этих деталей? Ответ (на бытовом уровне): средний, т.е. все размеры необходимо сложить и поделить на n — количество деталей. Но если величину $1/n$ посчитать "вероятностью соответствующего размера", то полученное среднее и есть математическое ожидание случайной величины X — измеряемый размер. Значение EX вычислено, как среднее значение наблюдаемых размеров, и не обязано совпадать с "ожиданием" мастера (например, если у последнего прибор имеет погрешность, то есть системную ошибку измерения). Итак, обнаруживается два разных "ожидания": ожидание мастера, изготовившего деталь, и "ожидание", вычисленное наблюдателем, которое по определению (7.1.2), и является математическим ожиданием.

Свойства математического ожидания.1) $Ec = c$.2) $E[cX] = cEX$.3) $E[X - EX] = 0$.

4) Для произвольных случайных величин X и Y , определенных на одном и том же вероятностном пространстве (Ω, \mathcal{U}, P) , имеет место равенство $E[X + Y] = EX + EY$.

5) Пусть X и Y — независимые случайные величины, определенные на одном и том же вероятностном пространстве (Ω, \mathcal{U}, P) . Тогда $EXY = EX \cdot EY$.

Доказательство. Свойства 1) и 2) с очевидностью вытекают из (7.1.1):

$$Ec = \int_{-\infty}^{+\infty} C p(x) dx = C \int_{-\infty}^{+\infty} p(x) dx = C.$$

$$E[cX] = \int_{-\infty}^{+\infty} C x p(x) dx = C \int_{-\infty}^{+\infty} x p(x) dx = cEX.$$

Поясним, однако, что постоянная c рассматривается как случайная величина, которая с вероятностью, равной 1, принимает значение c . Но тогда свойство 1) устанавливается и по формуле (7.1.2): $Ec = c \cdot 1 = c$.

Для доказательства свойства 3) обозначим $EX = m$ и введем случайную величину $Y = X - m$. Имеем:

$$EY = \int_{-\infty}^{+\infty} (x-m) p(x) dx = \int_{-\infty}^{+\infty} x p(x) dx - m \int_{-\infty}^{+\infty} p(x) dx = m - m = 0.$$

Замечание. Напомним, что значение EX имеет смысл центра распределения случайной величины. Центр распределения можно перенести в начало координат, введя случайную величину $X - EX$, поэтому она и называется *центрированной случайной величиной*.

Доказательство свойств 4) и 5) проведем в пункте 9.5.

7.2. Моменты. Дисперсия и среднее квадратическое отклонение

Понятие момента широко используется в механике (статические моменты, моменты инерции). Как уже отмечалось, статический момент мог быть интерпретирован как математическое ожидание. Но и моменты высших порядков используются в качестве числовых характеристик теории случайных величин.

Определение 7.2.1. *Начальным моментом α_r порядка r случайной величины X называется математическое ожидание случайной величины X^r :*

$$\alpha_r = EX^r = \int_{-\infty}^{\infty} x^r p(x) dx.$$

Для дискретной случайной величины, заданной законом $p(x) = \sum_{i=1}^n p_i \delta(x - x_i)$, момент порядка r определяется по формуле — $\alpha_r = \sum_i x_i^r p_i$.

Как видим, момент первого порядка случайной величины X совпадает с математическим ожиданием X .

На практике важны величины, характеризующие отклонение случайной величины от ее среднего значения (математического ожидания). Для этого применяются *центральные моменты порядка r* :

$$\mu_r = E[(X - EX)^r] = \int_{-\infty}^{\infty} (x - EX)^r p(x) dx.$$

Из свойства 3) математического ожидания видим, что центральные моменты первого порядка всегда равны нулю. Но тогда эти центральные моменты не могут быть характеристикой случайной величины. Поэтому следующей важнейшей характеристикой случайной величины является центральный момент вто-

рого порядка, называемый *дисперсией*¹⁰ случайной величины:

$$DX = E[(X - EX)^2] = \int_{-\infty}^{\infty} (x - EX)^2 p(x) dx.$$

Для дискретной случайной величины, заданной законом $p(x) = \sum_{i=1}^n p_i \delta(x - x_i)$, эта формула примет вид

$$DX = \sum_i (x_i - EX)^2 p_i.$$

Из определения дисперсии случайной величины видим, что это число имеет размерность, равную квадрату размерности случайной величины, поэтому удобнее и нагляднее характеристика, называемая *средним квадратическим отклонением*:

$$\sigma X = \sqrt{E[(X - EX)^2]}.$$

При вычислении дисперсии часто удобнее применить формулу

$$DX = EX^2 - E^2 X. \quad (7.2.1)$$

Докажем эту формулу. Имеем:

$$DX = E[X^2 - 2X EX + E^2 X].$$

Теперь воспользуемся свойствами математического ожидания 4) для суммы случайных величин и свойством 2) для произведения на константу. В результате получим (7.2.1):

$$DX = E[X^2 - 2EX EX + E^2 X] = EX^2 - E^2 X.$$

Свойства дисперсии.

- 1) $Dc = 0$.
- 2) $D[cX] = c^2 DX$.

¹⁰От латинского *dispersus* — рассеянный, рассыпанный.

3) Если X и Y — независимые случайные величины, определенные на вероятностном пространстве (Ω, \mathcal{U}, P) , то

$$D[X \pm Y] = DX + DY.$$

4) Если X и Y — независимые случайные величины, определенные на вероятностном пространстве (Ω, \mathcal{U}, P) , то

$$D[XY] = DX DY + DX E^2 Y + DY E^2 X.$$

Доказательства свойств 1) и 2) легко вытекают из соответствующих свойств математического ожидания, и мы их не приводим.

Доказательство свойства 3). Воспользовавшись определением дисперсии и свойствами 3) и 4) математического ожидания, имеем:

$$\begin{aligned} D[X \pm Y] &= E[(X \pm Y) - E[X \pm Y]]^2 = E[(X - EX) \pm (Y - EY)]^2 \\ &= E[(X - EX)^2 \pm 2E[(X - EX)(Y - EY)] + E[(Y - EY)^2]] = \\ &= DX \pm 2E[(X - EX)(Y - EY)] + DY. \end{aligned}$$

Среднее слагаемое исчезает, т.к. математическое ожидание центрированной случайной величины всегда равно нулю, и мы получаем утверждение свойства 3) дисперсии.

Доказательство свойства 4). Воспользовавшись определением дисперсии и свойствами 3) и 4) математического ожидания, имеем:

$$\begin{aligned} D[XY] &= E[XY - E[XY]]^2 = E[X^2 Y^2 - 2XYEXEY + E^2 X E^2 Y] = \\ &= EX^2 EY^2 - 2E^2 X E^2 Y + E^2 X E^2 Y = EX^2 EY^2 - E^2 X E^2 Y. \end{aligned}$$

Теперь, воспользовавшись равенством (7.2.1), можем записать $EX^2 = DX + E^2 X$ и $EY^2 = DY + E^2 Y$, откуда

$$\begin{aligned} D[XY] &= (DX + E^2 X)(DY + E^2 Y) - E^2 X E^2 Y = \\ &= DX DY + DX E^2 Y + DY E^2 X. \end{aligned}$$

7.3. Мода. Медиана

Определение 7.3.1. Модой случайной величины X (обозначение — $\mu(X)$) называют то ее значение, которому соответствует наибольшая вероятность или плотность вероятности.

Например, пусть дискретная случайная величина задана законом распределения

X	-2,7	-2	-1,3	0,7	1,7
P	0,15	0,35	0,4	0,12	0,08

Из таблицы видим, что самая большая вероятность, равная 0,4, отвечает значению случайной величины $X = -1,3$. Следовательно, $\mu(X) = -1,3$.

Определение 7.3.2. Медианой случайной величины X (обозначение — $me(X)$) называют такое ее значение $X = x$, при котором $P(X < x) = P(X > x)$, т.е. $S_1 = \int_{-\infty}^{x_0} p(x) dx = \int_{x_0}^{+\infty} p(x) dx = S_3$.

Ясно, что для симметричной случайной величины мода, медиана и математическое ожидание — одно и то же значение, являющееся центром симметрии. Но в общем случае эти значения (точки) не совпадают.

§ 8. Основные законы распределения

8.1. Равномерное распределение

Определение 8.1.1. Случайная величина называется распределенной равномерно на интервале (a, b) , если плотность распределения на этом интервале постоянна.

Предположим, что

$$p(x) = \begin{cases} 0, & x \notin (a, b) \\ C, & x \in [a, b]. \end{cases}$$

Поскольку

$$\int_{-\infty}^{\infty} p(x) dx = 1 = \int_{-\infty}^{\infty} C dx = C(b-a),$$

то $C = \frac{1}{b-a}$, следовательно,

$$p(x) = \begin{cases} 0, & x \notin (a, b) \\ \frac{1}{b-a}, & x \in [a, b]. \end{cases}$$

Найдем функцию распределения равномерно распределенной случайной величины.

1) Пусть $x \leq a$, тогда

$$F(x) = \int_{-\infty}^x p(t) dt = \int_{-\infty}^x 0 dt = 0.$$

2) Пусть $x \in [a, b]$, тогда

$$F(x) = \int_{-\infty}^x p(t) dt = \int_a^x \frac{1}{b-a} dt = \left. \frac{t}{b-a} \right|_a^x = \frac{x-a}{b-a}.$$

3) Пусть $x > b$, тогда

$$F(x) = \int_{-\infty}^x p(x) dx = \int_a^b \frac{1}{b-a} dx + \int_b^x 0 dx = 1.$$

Таким образом,

$$F(x) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & x \in [a, b]; \\ 1, & x > b. \end{cases}$$

Графики плотности $p(x)$ и функции $F(x)$ для равномерного распределения изображены на рисунке 12.

Найдем математическое ожидание равномерно распределенной случайной величины. Имеем:

$$EX = \int_a^b x p(x) dx = \int_a^b x \frac{1}{b-a} dx = \left. \frac{x^2}{2(b-a)} \right|_a^b = \frac{a+b}{2}.$$

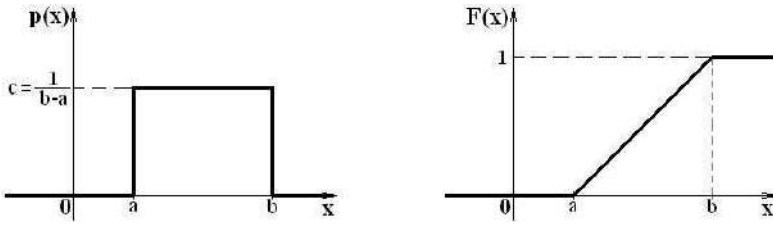


Рис. 12

Итак, математическое ожидание равномерно распределенной на отрезке $[a, b]$ случайной величины X совпадает с серединой отрезка $[a, b]$.

Теперь найдем дисперсию. Чтобы воспользоваться формулой (7.2.1), вычислим EX^2 . Имеем:

$$EX^2 = \frac{1}{b-a} \int_a^b x^2 dx = \frac{b^3 - a^3}{3(b-a)} = \frac{1}{3}(b^2 + ab + a^2).$$

Теперь, согласно формуле (7.2.1), получаем:

$$DX = \frac{1}{3}(b^2 + ab + a^2) - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

Для среднеквадратического отклонения получаем формулу:

$$\sigma X = \frac{b-a}{2\sqrt{3}}.$$

Отметим также, что в силу симметрии равномерно распределенной случайной величины относительно математического ожидания мода и медиана совпадают с математическим ожиданием.

Приведем формулу для определения вероятности попадания случайной величины в отрезок $[\alpha, \beta] \in [a, b]$:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} \frac{dx}{b-a} = \frac{\beta - \alpha}{b-a}.$$

8.2. Нормальное распределение

Вначале докажем равенство (интеграл Эйлера-Пуассона)

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1. \quad (8.2.1)$$

Для этого заметим, что

$$J = \int_{R_2} e^{-\frac{x^2+y^2}{2}} dx dy = \left[\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right]^2.$$

С другой стороны, интеграл J легко вычисляется в полярных координатах $x = r \cos \varphi$, $y = r \sin \varphi$, $r^2 = x^2 + y^2 \in [0, \infty)$, $\varphi \in [0, 2\pi]$:

$$\begin{aligned} J &= \int_0^{2\pi} d\varphi \int_0^{\infty} r e^{-\frac{r^2}{2}} dr = \\ &= \int_0^{2\pi} d\varphi \int_0^{\infty} e^{-\frac{r^2}{2}} d\left(\frac{r^2}{2}\right) = 2\pi. \end{aligned}$$

Отсюда следует (8.2.1).

Как мы знаем, любая функция p , для которой $\int_{R_1} p(x) dx = 1$, может служить функцией распределения некоторой случайной величины. Используя замену $\frac{x-a}{\sigma} = t$, а затем равенство (8.2.1) для произвольного a и $\sigma > 0$, получаем:

$$\frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = 1.$$

Определение 8.2.1. Случайная величина, заданная плотностью распределения

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad (8.2.2)$$

называется *нормальной* (часто говорят "распределенной по нормальному закону" или подчиняется закону распределения Гаусса) (рис.13).

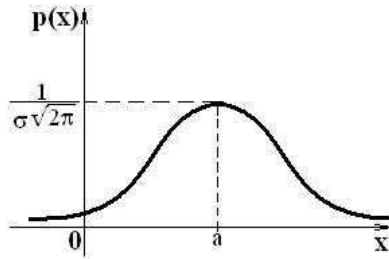


Рис. 13

Нормальное распределение относится к числу наиболее часто используемых на практике и теоретически изучаемых распределений. Применяется оно для приближенного описания многих случайных явлений, происходящих достаточно часто как в технических процессах, так и в психолого-педагогических, социальных, экономических и других реальных ситуациях, в которых на результат воздействует большое число факторов, среди которых нет явно выделяющихся. Например, при изучении рассеяния снарядов при стрельбе из артиллерийского орудия случайными факторами являются колебания воздуха, неодинаковость снарядов и др.

Нормальное распределение определяется двумя параметрами: a и σ . Достаточно знать эти параметры, чтобы задать нормальное распределение. На рисунке 14 наглядно показан геометрический смысл параметров a и σ . Прямая $x = a$ — ось симметрии кривой Гаусса. Большему значению σ соответствует более "крутая" кривая Гаусса. Определим вероятностный смысл этих параметров.

Математическое ожидание. $EX = a$.

Доказательство. Действительно, по определению математического ожидания имеем:

$$EX = \int_{-\infty}^{\infty} x p(x) dx = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

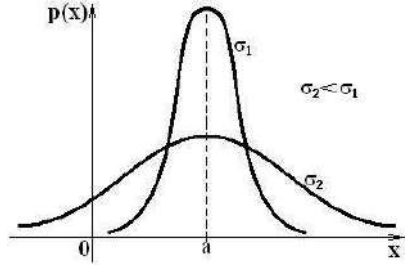


Рис. 14.

Произведем замену переменной $\frac{x-a}{\sigma} = t$, получим:

$$EX = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma t + a) e^{-\frac{t^2}{2}} dt =$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma t e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a e^{-\frac{t^2}{2}} dt.$$

Здесь первый интеграл от нечетной функции, понимаемый в смысле главного значения, и равен нулю, а второй вычисляется по формуле (8.2.1), и мы получаем: $EX = a$. Таким образом, если плотность определяется функцией $p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$, то $EX = a$.

Дисперсия. $DX = \sigma^2$.

Доказательство. Имеем:

$$DX = E[(X - EX)^2] = \int_{-\infty}^{\infty} (x - EX)^2 p(x) dx.$$

Подставим сюда функцию распределения (8.2.2) и произведем замену переменной $\frac{x-a}{\sigma} = t$, тогда

$$DX = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma t)^2 e^{-\frac{t^2}{2}} dt =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt \left(\frac{t^2}{2} \right).$$

Интегрируя по частям и полагая $u = t$,

$$v = \int e^{-\frac{t^2}{2}} d\left(\frac{t^2}{2}\right) = -e^{-\frac{t^2}{2}},$$

получим:

$$DX = \frac{\sigma^2}{\sqrt{2\pi}} \left(-te^{-\frac{t^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \right) = \sigma^2.$$

Таким образом,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad \implies \quad DX = \sigma^2.$$

Функция распределения и функция Лапласа. По определению функции распределения

$$F(x) = P\{X < x\} = \int_{-\infty}^x p(t) dt.$$

Для случайной величины, распределенной нормально, с плотностью (8.2.2), получим:

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt.$$

Произведем замену переменной $\frac{t-a}{\sigma} = y$, тогда

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^0 + \int_0^{\frac{x-a}{\sigma}} \right) e^{-\frac{y^2}{2}} dy.$$

Первый интеграл вычисляется по формуле (8.2.1) и равен $\frac{\sqrt{2\pi}}{2}$, а второй — неопределенный интеграл называется *функцией Лапласа*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{y^2}{2}} dy.$$

Таким образом,

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right).$$

Как известно, вероятность попадания случайной величины в интервал (α, β) вычисляется по формуле

$$P\{\alpha < X < \beta\} = F(\beta) - F(\alpha).$$

Отсюда

$$P\{\alpha < X < \beta\} = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \quad (8.2.3)$$

Правило трех сигм. Случайная величина X , распределенная по нормальному закону с дисперсией σ^2 , отклоняется от своего математического ожидания $EX = a$ на расстояние больше чем 3σ крайне редко.

Более точно:

$$P\{|X - a| < 3\sigma\} \approx 0,9973. \quad (8.2.4)$$

Это практическое правило, т.е. если Вы столкнулись с такой случайной величиной, то можете утверждать, что все практические реализации ее находятся в области $\{X = x \in [a - 3\sigma, a + 3\sigma]\}$.

Вначале докажем более общее утверждение:

$$P\{|X - a| < \varepsilon\} = 2\Phi\left(\frac{\varepsilon}{\sigma}\right) \quad (8.2.5)$$

Для доказательства (8.2.5) заметим, что

$$\begin{aligned} |X - a| < \varepsilon &\iff a - \varepsilon < X < a + \varepsilon, \implies \\ \implies P\{|X - a| < \varepsilon\} &= P\{a - \varepsilon < X < a + \varepsilon\}. \end{aligned}$$

Теперь воспользуемся равенством (8.2.3) и получим (8.2.5).

Полагая в (8.2.5) $\varepsilon = t\sigma$, получим:

$$P\{|X - a| < t\sigma\} = 2\Phi(t). \quad (8.2.6)$$

По таблице функций Лапласа (см. например [5]) находим:

$$\begin{aligned} \text{при } t = 1 \quad P\{|X - a| < \sigma\} &= 2\Phi(1) \approx 0,6827; \\ \text{при } t = 2 \quad P\{|X - a| < 2\sigma\} &= 2\Phi(2) \approx 0,9545, \\ \text{при } t = 3 \quad P\{|X - a| < 3\sigma\} &= 2\Phi(3) \approx 0,9973. \end{aligned}$$

откуда и вытекает правило трех сигм (8.2.4).

8.3. Показательное распределение

Определение 8.3.1. *Непрерывную случайную величину, плотность распределения которой определяется выражением*

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0, \end{cases} \quad (8.3.1)$$

называют случайной величиной, распределенной по показательному (экспоненциальному) закону.

Это распределение часто наблюдается при изучении сроков службы различных устройств (электронных, механических), времени безотказной работы различных систем или элементов систем. Параметр λ обычно называют *интенсивностью отказов*.

Функция распределения определяется интегрированием (8.3.1):

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases} \quad (8.3.2)$$

Графики плотности $p(x)$ и функции $F(x)$ показательного распределения приведены на рисунке 15.

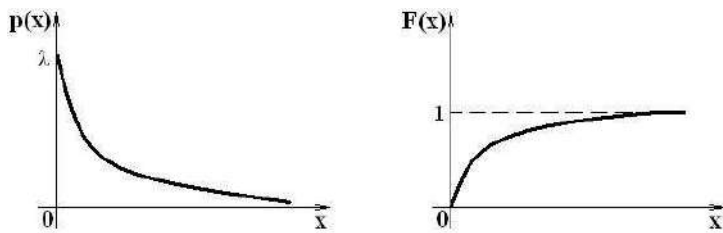


Рис. 15

Интегрированием по частям легко определяются математическое ожидание и дисперсия:

$$EX = \frac{1}{\lambda}, \quad DX = \frac{1}{\lambda^2}.$$

Соответственно среднее квадратическое отклонение $\sigma = \frac{1}{\lambda}$.

8.4. Биномиальное распределение

Хорошо известна формула Бинома (называется также "формула бинома Ньютона")

$$(p + q)^n = \sum_{k=0}^n C_n^k p^k q^{n-k}, \quad C_n^k = \frac{n!}{k!(n-k)!}.$$

Полагая здесь $p > 0$, $q > 0$, $p + q = 1$, получим:

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

Но тогда число $C_n^k p^k q^{n-k}$ представляет вероятность того, что в схеме испытаний Бернулли событие, появляющееся в каждом испытании с вероятностью p и не происходящее с вероятностью $q = 1 - p$, в n испытаниях появится ровно k раз. Поэтому предыдущая формула примет вид

$$1 = \sum_{k=0}^n P_n(k).$$

Определение 8.4.1. *Дискретная случайная величина, заданная законом распределения*

X	x_1	x_2	x_3	\dots	x_n
P	$p_n(1)$	$p_n(2)$	$p_n(3)$	\dots	$p_n(n)$

где вероятности $p_n(k)$ определены по формуле Бернулли:

$$p_n(k) = P\{X = k\} = C_n^k p^k q^{n-k}, \quad C_n^k = \frac{n!}{k!(n-k)!},$$

называется *распределенной биномиально*. Вид биномиального распределения для различных значений p представлен на рисунке 16.

Биномиальное распределение — это одно из самых распространенных дискретных распределений. Чаще всего встречается

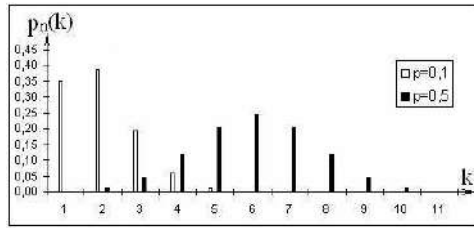


Рис. 16

на практике, например, при изучении таких явлений, как статистический контроль качества продукции, поскольку любой технологический процесс производства математически представляет собой схему испытаний Бернулли; число людей в выборке, имеющих данную группу крови, и т.д.

Без доказательства приведем формулы для математического ожидания, дисперсии и среднеквадратического отклонения:

$$EX = np, \quad DX = npq = np(1 - p), \quad \sigma X = \sqrt{npq}.$$

Отметим также, что, как это следует из локальной и интегральной теоремы Лапласа, биномиальное распределение приближается к нормальному с ростом n или при малых значениях p или q .

8.5. Распределение Пуассона

Распределение Пуассона — это дискретное распределение, являющееся одним из важных предельных случаев биномиального распределения, когда вероятность наступления интересующего события в единичном испытании очень мала ($p \rightarrow 0$), но число испытаний, производимых в единицу времени, достаточно велико ($n \rightarrow \infty$). Таким образом, при росте n и фиксированном значении произведения $np = \lambda > 0$ биномиальное распределение сводится к распределению Пуассона.

Определение 8.5.1. *Распределением Пуассона называется распределение вероятностей между целыми неотрицательны-*

ми числами, когда на долю числа m приходится вероятность

$$P(m, \lambda) = \frac{\lambda^m}{m!} e^{-\lambda}, \quad \lambda > 0, m = 0, 1, 2, \dots; \quad (8.5.1)$$

Формула (8.5.1) задает распределение вероятностей на множестве целых неотрицательных чисел, так как $P(m, \lambda) > 0$ для всех $m = 0, 1, \dots$ и $\sum_{m=0}^{\infty} P(m, \lambda) = 1$.

Доказательство последнего равенства опирается на разложение в ряд Маклорена функции e^λ :

$$\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^\lambda$$

для любого $\lambda > 0$.

Определение 8.5.2. Дискретная случайная величина, заданная законом распределения

X	x_1	x_2	\dots	x_m	\dots
P	$p(1)$	$p(2)$	\dots	$p(m)$	\dots

где вероятности $p(m)$ определены по формуле:

$$p(m) = P\{X = m\} = \frac{\lambda^m}{m!} e^{-\lambda}, \quad \lambda > 0, m = 0, 1, 2, \dots,$$

называется *распределенной по Пуассону с параметром $\lambda > 0$* . Вид распределения Пуассона для различных значений $\lambda > 0$ представлен на рисунке 17.

Читателю предлагаем самостоятельно с использованием указанного выше разложения доказать формулы математического ожидания, дисперсии и среднеквадратического отклонения случайной величины, распределенной по Пуассону:

$$EX = \lambda, \quad DX = \lambda, \quad \sigma X = \sqrt{\lambda}.$$

Параметр λ пуассоновского распределения равен одновременно математическому ожиданию и дисперсии случайной величины X , имеющей изучаемое распределение. В этом состоит особенность данного распределения, которая используется на практике.

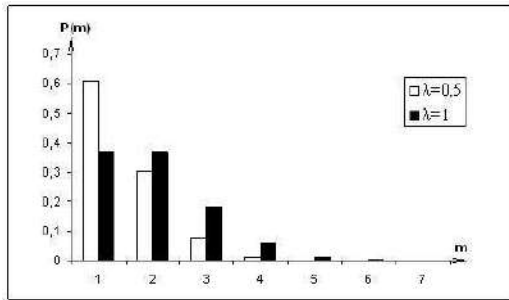


Рис. 17

Распределение Пуассона играет важную роль в приложениях теории вероятностей и математической статистики, в теории массового обслуживания, страховании, теории надежности, теории связи, ядерной физике или физике элементарных частиц и т.п. — всюду, где в течение определенного времени может происходить случайное число редких событий. Хорошо описывается формулой Пуассона процесс радиоактивного распада, например, радия. Этот процесс заключается в превращении ядра атома радия в ядро атома радона с испусканием альфа-частицы. Распад каждого отдельного ядра происходит независимо от состояния других ядер, и вероятность такого распада в единицу времени есть величина постоянная. Примерами случайных величин, распределенных по Пуассону в технике, являются число сбоев на автоматической линии; число отказов технического устройства высокой надежности, работающего в нормальном режиме, при многократном применении; число поврежденных ячеек в LCD-матрице, образующихся при долговременном интенсивном использовании ноутбука и др.

8.6. Геометрическое распределение

Геометрическое распределение в теории вероятностей — распределение дискретной случайной величины X , равной количеству испытаний в схеме Бернулли до наблюдения первого «успеха». Основными параметрами геометрического распределе-

ния являются число m – число «неудач» до первого «успеха», $m = 0, 1, 2, 3, \dots$; p – вероятность «успеха»; $q = 1 - p$ – вероятность «неудачи».

Определение 8.6.1. *Дискретная случайная величина, заданная законом распределения*

X	x_1	x_2	\dots	x_m	\dots
P	$p(1)$	$p(2)$	\dots	$p(m)$	\dots

где вероятности $p(m)$ определены по формуле:

$$p(m) = P\{X = m\} = pq^m,$$

$0 < p < 1$; $q = 1 - p$; $m = 0, 1, 2, \dots$, называется *распределенной по геометрическому закону*.

Вероятности $p(m)$ образуют геометрическую прогрессию с первым членом p и знаменателем q : p, qp, q^2p, q^3p, \dots . По этой причине распределение называют геометрическим. Легко убедиться, что ряд сходится и сумма его равна единице. Действительно, сумма ряда есть сумма членов бесконечной геометрической прогрессии со знаменателем меньшим единицы, тогда сумма его:

$$p + qp + q^2p + q^3p + \dots = \frac{p}{1 - q} = \frac{p}{p} = 1.$$

Найдем числовые характеристики случайной величины X , распределенной по геометрическому закону. С этой целью воспользуемся аппаратом производящих функций.

Определение 8.6.2. *Производящей функцией для случайной величины X называется функция вида*

$$\varphi(z) = \sum_{k=0}^{\infty} p_k z^k,$$

где z – произвольный параметр ($0 < z \leq 1$).

Беря первую производную по z от производящей функции и положив в ней $z = 1$, имеем:

$$\varphi'(z) = \sum_{k=0}^{\infty} k p_k z^{k-1},$$

$$\varphi'(1) = \sum_{k=0}^{\infty} k p_k = EX.$$

Таким образом, математическое ожидание неотрицательной целочисленной случайной величины равно первой производной ее производящей функции $\varphi(z)$ при $z = 1$.

Аналогично можно установить, что дисперсия случайной величины определяется равенством:

$$DX = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2.$$

Математическое ожидание. $EX = \frac{q}{p}$.

Доказательство. Действительно, запишем производящую функцию для случайной величины X , распределенной по геометрическому закону:

$$\varphi(z) = \sum_{m=0}^{\infty} p(qz)^m = p \sum_{m=0}^{\infty} (qz)^m = \frac{p}{1 - qz},$$

т.е.

$$\varphi(z) = \frac{p}{1 - qz}.$$

Далее последнее выражение дифференцируем по z , имеем:

$$\varphi'(z) = \frac{pq}{(1 - qz)^2}.$$

Отсюда находим математическое ожидание:

$$EX = \varphi'(1) = \frac{pq}{(1 - q)^2} = \frac{pq}{p^2} = \frac{q}{p}.$$

Дисперсия. $EX = \frac{q}{p^2}$.

Доказательство. Беря вторую производную по z от производящей функции для случайной величины X , распределенной по геометрическому закону, имеем:

$$\varphi''(z) = \frac{2pq^2}{(1 - qz)^3} = \frac{2pq^2}{p^3} = \frac{2q^2}{p^2}.$$

Далее находим дисперсию изучаемой случайной величины X :

$$DX = \frac{2q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{2q^2 + qp - q^2}{p^2} = \frac{q^2 + qp}{p^2} = \frac{q(q+p)}{p^2} = \frac{q}{p^2}.$$

И, значит, $\sigma X = \frac{\sqrt{q}}{p}$.

Примерами случайных величин, имеющих геометрическое распределение, являются следующие: число испытаний прибора до первого отказа; число циклов обзора радиолокатора до обнаружения объекта; количество включений телевизионного приемника до его первого выхода из строя; число запусков жесткого диска до образования первых повреждений физической поверхности нулевого трека; число подключений Flash-памяти к USB порту до первого повреждения прошивки интерфейсного контроллера и др.

§ 9. Системы случайных величин

9.1. Понятие системы случайных величин

Ясно, что на одном и том же вероятностном пространстве (Ω, \mathcal{U}, P) можно задать несколько случайных величин X_1, X_2, \dots, X_n . Тем самым на пространстве (Ω, \mathcal{U}, P) определен случайный вектор $\mathbf{X} = (X_1, X_2, \dots, X_n)$, возможные значения которого (или реализация случайной величины) есть n -мерный вектор $\hat{x} = (x_1, x_2, \dots, x_n)$. Удобно интерпретировать этот вектор в виде точки в n -мерном евклидовом пространстве точек R_n .

Определение 9.1.1. *Несколько случайных величин $\mathbf{X} = (X_1, X_2, \dots, X_n)$, распределенных совместно на одном и том же вероятностном пространстве (Ω, \mathcal{U}, P) , называется системой случайных величин, размерности n . Обозначение $\mathbf{X}[n]$.*

♦ **Пример 1.** Станок-автомат штампует стальные плитки. Пространство элементарных событий $\Omega = \{A_1 = \text{"плитка стандартная"}, A_2 = \text{"плитка нестандартная"}\}$. Роль случайных вели-

чин могут выполнить размеры плитки: X_1 — длина, X_2 — ширина, X_3 — высота. Таким образом, определена 3-х мерная система случайных величин $\mathbf{X}[3]$. Если контролируется еще и масса плитки $m = X_4$, мы получим 4-х мерную систему случайных величин $\mathbf{X}[4] = (X_1, X_2, X_3, X_4)$.

Определение 9.1.2. *Законом распределения системы случайных величин называется любое соотношение, устанавливающее связь между возможными значениями случайных величин и их вероятностями.*

Так же, как и в одномерном распределении, закон распределения может быть задан в различных формах. В дискретном случае в виде "многомерных" таблиц (см. таблицу 7.1 для $\mathbf{X}[2] = (X, Y)$), функций распределения и плотности распределения систем случайных величин.

Определение 9.1.3. *Функцией распределения системы случайных величин называется вероятность совместного выполнения неравенств $X_1 < x_1, X_2 < x_2, \dots, X_n < x_n$:*

$$F(x_1, x_2, \dots, x_n) = P\{X_1 < x_1, X_2 < x_2, \dots, X_n < x_n\}.$$

Определение 9.1.4. *Пусть функция распределения F системы случайных величин $\mathbf{X}[n]$ в некоторой области $D \in R^n$ является n раз непрерывно дифференцируемых функций, т.е. $F \in C_n(D)$. Плотностью вероятности системы $\mathbf{X}[n]$ называется смешанная производная порядка n от функции распределения:*

$$p(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}.$$

Вероятностный смысл этой функции заключается в следующем. Если через dx_1, dx_2, \dots, dx_n обозначить дифференциалы независимых переменных и придать им смысл размеров некоторого n -мерного параллелепипеда $\Delta B \in D$, одна из вершин которого — точка $x = (x_1, x_2, \dots, x_n)$, то

$$P\{\mathbf{X}[n] \in B\} = p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

Подробнее мы рассмотрим эти понятия на примере 2-х мерных случайных величин.

9.2. Функция распределения системы $X[2]$

Простейшая система $\mathbf{X}[2] = \{X, Y\}$ состоит из двух случайных величин X и Y . Ее можно рассматривать как координаты случайной точки (x, y) на плоскости с декартовой системой координат xOy .

Определение 9.2.1. *Функцией распределения системы двух случайных величин X и Y называется вероятность совместного выполнения неравенств $X < x, Y < y$:*

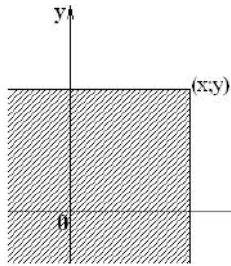


Рис. 18

$$F(x, y) = P\{X < x, Y < y\}.$$

Геометрическая интерпретация: вероятность того, что случайная точка на плоскости xOy попадет в квадрант плоскости с вершиной в точке (x, y) , лежащий левее и ниже ее (рис. 18).

Вероятность того, что случайная точка на плоскости попадет в какую-либо полуплоскость (по известной функции распределения F) находится по формулам, соответственно

$$\begin{aligned} P\{X[2] = (X, Y) : X < x\} &= \lim_{y \rightarrow \infty} F(x, y) = F(x, \infty) = F_1(x), \\ P\{X[2] = (X, Y) : Y < y\} &= \lim_{x \rightarrow \infty} F(x, y) = F(\infty, y) = F_2(y). \end{aligned} \quad (9.2.1)$$

Вероятность того, что случайная точка на плоскости попадет в какую-либо точку прямоугольника

$$\{X[2] = (X, Y) : (x_1 < X < x_2, y_1 < Y < y_2)\}$$

(по известной функции распределения F) определяется по формуле

$$\begin{aligned} P\{x_1 < X < x_2, y_1 < Y < y_2\} &= F(x_2, y_2) - F(x_2, y_1) - \\ &\quad - F(x_1, y_2) + F(x_1, y_1). \end{aligned} \quad (9.2.2)$$

Доказательство равенства (9.2.2). Множество точек прямоугольника $A_1(x_1, y_1), A_2(x_2, y_1), A_3(x_2, y_2), A_4(x_1, y_2)$ обозначим через B (рис. 19). Множество точек квадранта с вершиной в точке (x_i, y_j) удобно обозначить $\{x_i, y_j\}$. Так, что $P\{x_i, y_j\}$ — вероятность, что точка $(X, Y) \in \{x_i, y_j\}$. Имеем:

$$\{x_2, y_2\} = B + \{x_2, y_1\} + \{x_1, y_2\}.$$

Согласно аксиоме АЗ

$$P\{x_2, y_2\} = P\{B\} + P(\{x_2, y_1\} + \{x_1, y_2\}).$$

Второе слагаемое найдем по теореме сложения

$$P(\{x_2, y_1\} + \{x_1, y_2\}) = P\{x_2, y_1\} + P\{x_1, y_2\} - P\{x_1, y_1\}.$$

Учитывая, что $P\{x_i, y_j\} = F(x_i, y_j)$, получаем утверждение (9.2.2).

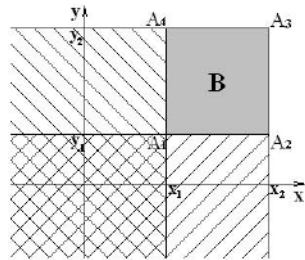


Рис. 19

Свойства функции распределения $F(x, y)$.

1. $F(x, y)$ является неубывающей по каждому из аргументов.
2. $F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0$.
3. $F(+\infty, +\infty) = 1$.

Доказательства свойств 1-3 вытекают непосредственно из определения функции распределения, и мы их опустим.

9.3. Плотность вероятности системы $X[2]$

Пусть в области D , принадлежащей плоскости xOy , функция распределения вероятности некоторой случайной величины представляет собой дважды непрерывно дифференцируемую функцию $F(x, y)$, т.е. $F(x, y) \in C_2(D)$. **Средней плотностью вероятности попадания случайной величины в некоторый прямоугольник ΔD с размерами Δx и Δy , принадлежащими области D , является отношение вероятности попадания случайной величины $X[2]$ в ΔD к площади этого прямоугольника:** $\frac{P\{X[2] \in \Delta D\}}{\Delta x \Delta y}$. Плотностью распределения случайной величины в точке естественно назвать предел (если он существует)

$$p(x, y) = \lim_{\Delta y \rightarrow 0} \left(\lim_{\Delta x \rightarrow 0} \frac{P\{X[2] \in \Delta D_j\}}{\Delta x \Delta y} \right).$$

Существование этого предела в данном случае обеспечено тем, что $F(x, y) \in C^2(D)$. В самом деле, согласно свойству 5 функции $F(x, y)$ имеем:

$$\begin{aligned} P\{X \in [x, x + \Delta x], Y \in [y, y + \Delta y]\} &= \\ &= F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - (F(x + \Delta x, y) - F(x, y)). \end{aligned}$$

Применяя формулу Лагранжа "конечных приращений" получим:

$$\begin{aligned} P\{X \in [x, x + \Delta x], Y \in [y, y + \Delta y]\} &= F'_x(\xi_1, y + \Delta y) \Delta x - \\ &- F'_x(\xi_2, y + \Delta y) \Delta x, \end{aligned}$$

где ξ_1 и ξ_2 — соответствующие средние точки из интервала $(x, x + \Delta x)$. Ясно, что

$$\left. \begin{array}{c} \xi_1 \\ \xi_2 \end{array} \right\} \rightarrow x$$

при $\Delta x \rightarrow 0$. Поэтому

$$\lim_{\Delta x \rightarrow 0} \frac{P\{X[2] \in \Delta D_j\}}{\Delta x \Delta y} = \frac{1}{\Delta y} [F'_x(x, y + \Delta y) - F'_x(x, y)].$$

Снова применяя формулу Лагранжа (по y) и переходя к пределу при $\Delta y \rightarrow 0$, получим равенство

$$p(x, y) = \frac{\partial^2 F(x, y)}{\partial y \partial x},$$

которое примем за определение плотности распределения вероятности случайной величины $X[2]$ по известной функции распределения вероятности $F(x, y)$ этой случайной величины.

Если плотность на некотором прямоугольнике D постоянна $p(x, y) = p_0$, то из определения средней плотности имеем:

$$P\{X[2] \in D\} = p_0 \Delta x \Delta y,$$

где Δx и Δy стороны прямоугольника D .

Предположим, что нам известна плотность $p(x, y)$ распределения некоторой случайной величины и что

(i) *функция $p(x, y)$ интегрируема по Риману в области D .*

Тогда для приближенного вычисления вероятности попадания случайной величины $X[2]$ в область D можно применить следующий процесс: область D разбиваем на достаточно мелкие части ΔD_j , в каждой из которых фиксируется произвольная точка (x_j, y_j) , а плотность на этом участке разбиения приближенно считается постоянной, равной $p(x_j, y_j)$, и тогда

$$P\{X[2] \in D\} = \sum_{n=1}^n P\{X[2] \in \Delta D_j\} \approx \sum_{n=1}^n p(x_j, y_j) |\Delta D_j|.$$

Здесь последнее выражение представляет собой сумму Римана для двойного интеграла по области D , который существует в виду условия (i), и поэтому точное значение искомой вероятности находится по формуле

$$P\{X[2] \in D\} = \int_D p(x, y) dx dy.$$

Теперь необходимо отметить следующее. Мы определили плотность распределения случайной величины $X[2]$ как смешанную производную от функции распределения $F(x, y)$ при условии, что $F \in C_2(D)$. Это условие может быть нарушено. Например, функция распределения представляет собой (негладкую) склейку нескольких функций. Но тогда и функция плотности $p(x, y)$ представляет собой также склейку соответствующих функций и не может быть получена (в произвольной точке) как смешанная производная от функции распределения. Более общим является следующее определение плотности распределения вероятности случайной величины $X[2]$.

Определение 9.3.1. Функция $p(x, y)$ называется плотностью распределения вероятности случайной величины $X[2]$, если вероятность попадания случайной величины в произвольную область G плоскости xOy вычисляется по формуле

$$P\{X[2] \in G\} = \int_G p(x, y) dx dy.$$

В этом определении функция p как функция точки (x, y) , определена почти всюду, т.е. с точностью до множества точек с нулевой лебеговой мерой.

Свойства плотности распределения случайной величины функции

Установите самостоятельно следующие свойства функции p :

1. $p(x, y) \geq 0$;
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1$;
3. $\int_{-\infty}^x \int_{-\infty}^y p(x, y) dx dy = F(x, y)$;
4. $\int_{-\infty}^x \int_{-\infty}^{\infty} p(x, y) dx dy = F_1(x)$; $\int_{-\infty}^{\infty} \int_{-\infty}^y p(x, y) dx dy = F_2(y)$;
5. $\int_{-\infty}^{\infty} p(x, y) dy = p_1(x)$; $\int_{-\infty}^{\infty} p(x, y) dx = p_2(y)$;
6. $F_1'(x) = p_1(x)$; $F_2'(y) = p_2(y)$.

Следует обратить внимание, что функции F_1 и F_2 , определенные в равенствах (9.2.1) и определенные в свойстве 4, имеют один и тот же вероятностный смысл (вероятность попасть в полуплоскость), поэтому совпадают.

9.4. Зависимые и независимые случайные величины

Определение 9.4.1. *Условным законом распределения одной случайной величины, входящей в систему, называется закон, найденный при условии, что другая случайная величина из этой же системы приняла определенное значение.*

Из определения следует, что условная функция вероятности и условная плотность вероятности равны соответственно:

$$F(x/y) = P\{X < x/Y = y\}, \quad p(x/Y = y) = \frac{dF(x/Y = y)}{dx}. \quad (9.4.1)$$

Если функция распределения $F(x/y)$ дифференцируема, то условная плотность распределения вероятности определяется равенством

$$p(x/Y = y) = \frac{dF(x/Y = y)}{dx}.$$

Случайные величины X , Y , определенные на одном и том же вероятностном пространстве (Ω, \mathcal{U}, P) , называются независимыми, если их функция совместного распределения представлена в виде произведения функций, зависящих только от x и только от y :

$$F(x, y) = F_1(x) F_2(y). \quad (9.4.2)$$

Теорема 9.4.1. *Если X и Y случайные величины, определенные на вероятностном пространстве (Ω, \mathcal{U}, P) , независимы, то*

$$P\{x_1 < X < x_2, y_1 < Y < y_2\} = P\{x_1 < X < x_2\} P\{y_1 < Y < y_2\}.$$

Доказательство. Используя формулу (9.2.2) и определение (9.4.2.) вероятности попадания случайной величины $X[2]$ в прямоугольник, получаем:

$$P\{x_1 < X < x_2, y_1 < Y < y_2\} = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) +$$

$$\begin{aligned}
& +F(x_1, y_1) = F_1(x_2) F_2(y_2) - F_1(x_1) F_2(y_2) - F_1(x_2) F_2(y_1) + \\
& +F_1(x_1) F_2(y_1) = F_2(y_2) [F_1(x_2) - F_1(x_1)] + F_2(y_1) [F_1(x_2) - F_1(x_1)] = \\
& = [F_1(x_2) - F_1(x_1)] [F_2(y_2) - F_2(y_1)] = \\
& = P\{x_1 < X < x_2\} P\{y_1 < Y < y_2\}.
\end{aligned}$$

Доказательство закончено.

Теорема 9.4.2. Для того чтобы случайные величины X и Y , определенные на вероятностном пространстве (Ω, \mathcal{U}, P) , были независимы, необходимо и достаточно, чтобы их совместная плотность распределения вероятности представлялась в виде произведения функций, зависящих только от x и только от y :

$$p(x, y) = p_1(x) p_2(y). \quad (9.4.3)$$

Доказательство. 1. Н е о б х о д и м о с т ь. Если X и Y — независимые случайные величины, то согласно теореме 9.4.1

$$P\{x_1 < X < x_2, y_1 < Y < y_2\} = P\{x_1 < X < x_2\} P\{y_1 < Y < y_2\}.$$

Отсюда следует равенство

$$\int_D (x, y) dx dy = \int_{x_1}^{x_2} p_1(x) dx \int_{y_1}^{y_2} p_2(y) dy,$$

где $D = \{(x, y) : x_1 < X < x_2, y_1 < Y < y_2\}$. Поскольку вершины этого прямоугольника произвольны, а функции, входящие в полученное равенство непрерывны, то отсюда следует равенство (9.4.3).

2. Д о с т а т о ч н о с т ь. Пусть выполняется равенство

$$p(x, y) = p_1(x) p_2(y).$$

Интегрируя его по квадранту $\{(x, y) : X \in (-\infty, x), Y \in (-\infty, y)\}$, получим (9.4.2.):

$$F(x, y) = F_1(x) F_2(y),$$

отсюда независимость случайных величин X и Y следует по определению. Доказательство закончено.

9.5. Моменты, математическое ожидание, дисперсия системы $X[2]$

Пусть $p(x, y)$ — плотность распределения случайной величины $X[2] = (X, Y)$. Как нам уже известно, моментами случайной величины (одномерной) называется математическое ожидание степени случайной величины. Для $X[2]$ моментом порядка (α, β) называется математическое ожидание произведения степеней случайных величин:

$$[X^\alpha, Y^\beta].$$

Выделим моменты порядков $(1, 0)$, $(0, 1)$ и $(2, 0)$ и $(0, 2)$ (это соответственно математическое ожидание и дисперсия).

Определение 9.5.1 Математическим ожиданием случайных величин X и Y , входящих в систему $X[2]$, называются следующие значения

$$EX = \sum_{i,j} x_i p(x_i y_j), \quad EX = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x p(x, y) dx dy = \int_{-\infty}^{+\infty} x p_1(x) dx;$$

$$EY = \sum_{i,j} y_j p(x_i y_j), \quad EY = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y p(x, y) dx dy = \int_{-\infty}^{+\infty} y p_2(y) dy$$

для дискретных и непрерывных случайных величин соответственно. Здесь введены плотности распределений $p_1(x)$ и $p_2(y)$ случайных величин X и Y , входящих в систему $X[2]$:

$$p_1(x) = \int_{-\infty}^{+\infty} p(x, y) dy, \quad p_2(y) = \int_{-\infty}^{+\infty} p(x, y) dx.$$

Докажем свойства математического ожидания 4) и 5), сформулированные в пункте 7.1.

Д о к а з а т е л ь с т в о свойства 4): $E[X + Y] = EX + EY$.

Достаточно доказать свойство для непрерывных случайных величин. Поскольку случайные величины X и Y распределены

на одном вероятностном пространстве (Ω, \mathcal{U}, P) , то и случайная величина $X + Y$ распределена на этом же вероятностном пространстве, поэтому

$$\begin{aligned} E[X + Y] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) p(x, y) dx dy = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x p(x, y) dx dy + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y p(x, y) dx dy. \end{aligned}$$

Исходя из определения математического ожидания для случайных величин, входящих в систему $X[2]$, получаем свойство 4).

Д о к а з а т е л ь с т в о 5): $E[XY] = EX \cdot EY$. Достаточно доказать свойство для непрерывных случайных величин. Поскольку случайные величины X и Y распределены на одном вероятностном пространстве (Ω, \mathcal{U}, P) , то и случайная величина $X + Y$ распределена на этом же вероятностном пространстве, поэтому

$$E[XY] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x y p(x, y) dx dy.$$

Для независимых случайных величин X и Y плотность распределения $p(x, y) = p_1(x) p_2(y)$. Поэтому

$$E[XY] = \int_{-\infty}^{+\infty} x p_1(x) \int_{-\infty}^{+\infty} y p_2(y) dx dy = EY \cdot EX.$$

Таким образом, для независимых случайных величин X, Y , распределенных на одном вероятностном пространстве (Ω, \mathcal{U}, P) , получаем свойство 5): $E[XY] = EX EY$.

Пусть $X[2] = (X, Y)$. Введем обозначение для центрированных случайных величин, соответственно $X_c = X - EX$, $Y_c = Y - EY$.

Определение 9.5.2. Дисперсией случайных величин X и Y , входящих в систему $X[2]$, называются следующие значения

$$DX = E[X_c^2 Y_c^0] = E[X_c^2], \quad DY = E[X_c^0 Y_c^2] = E[Y_c^2].$$

В пункте 7.2 для независимых случайных величин X и Y , входящих в систему $X[2]$, мы доказали свойство $D[X + Y] = DX + DY$. Для произвольных $(X, Y) = X[2]$ имеем (см. параграф 7.2, доказательство свойства 3 дисперсии)

$$D[X + Y] = DX + DY + 2E[(X - EX)(Y - EY)].$$

Причем здесь третье слагаемое равно нулю для независимых случайных величин X и Y , поэтому неравенство нулю выражения

$$E[(X - EX)(Y - EY)] \neq 0$$

есть признак наличия зависимости между X и Y . Математическое ожидание произведения центрированных случайных величин $(X - EX)$ и $(Y - EY)$ называется *корреляционным моментом* $E[X_c Y_c]$ (*моментом связи*) или *ковариацией*¹¹ случайных величин X, Y , входящих в систему $X[2]$.

9.6. Ковариация случайных величин. Коэффициент корреляции

Пусть случайные величины X и Y определены на вероятностном пространстве (Ω, \mathcal{U}, P) . Ковариацией случайных величин X и Y называется математическое ожидание произведения $(X - EX)(Y - EY)$:

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)].$$

Свойства

$$1) \quad \text{cov}(X, Y) = \text{cov}(Y, X);$$

¹¹От английского covariation - совместная вариация, абсолютный показатель совместного рассеяния значений двух случайных величин. Может рассматриваться как двумерный аналог дисперсии.

- 2) $\text{cov}(X, X) = DX$;
 3) ковариация независимых случайных величин равна нулю;
 4) $D[\sum_{i=1}^n c_i X_i] = \sum_{i,j=1}^n c_i c_j \text{cov}(X_i, X_j)$;
 5) $|\text{cov}(X, Y)| \leq \sqrt{DX DY}$.

Доказательство. Свойства 1), 2) и 3) с очевидностью вытекают из определения функции cov .

Доказательство свойства 4). Рассмотрим случайную величину $Y = \sum_{i=1}^n c_i \cdot X_i$. Имеем: $EY = \sum_{i=1}^n c_i \cdot EX_i$. Тогда

$$\begin{aligned} D\left[\sum_{i=1}^n c_i X_i\right] &= E[(Y - EY)^2] = \\ &= E\left[\sum_{i=1}^n c_i (X_i - EX_i) \sum_{j=1}^n c_j (X_j - EX_j)\right] = \\ &= E\sum_{i,j=1}^n c_i c_j (X_j - EX_j)(X_i - EX_i) = \\ &= \sum_{i,j=1}^n c_i c_j E[(X_j - EX_j)(X_i - EX_i)] = \sum_{i,j=1}^n c_i c_j \text{cov}(X_i, X_j). \end{aligned}$$

Доказательство свойства 5). Применяя свойства дисперсии и свойства 3) и 2), получим:

$$0 \leq D[c_1 X + c_2 Y]^2 = c_1^2 DX + c_2^2 DY + 2c_1 c_2 \text{cov}(X, Y).$$

Выражение справа рассматриваем как квадратичную форму от c_1, c_2 . Знакоопределенность квадратической формы говорит о том, что дискриминант квадратного трехчлена (относительно c_1 или c_2) не положителен, т.е. $c_2^2 \text{cov}^2(X, Y) - c_2^2 DX DY \leq 0$, откуда следует утверждение свойства 4).

Свойство 4) дает основание рассматривать в качестве количественной характеристики зависимости случайных величин сле-

дующее отношение, которое называется **коэффициентом корреляции**¹² случайных величин X и Y :

$$K(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{DX DY}}.$$

Свойства коэффициента корреляции.

- 1) $|K(X, Y)| \leq 1$;
- 2) если X и Y — независимые случайные величины, то $K(X, Y) = 0$;
- 3) если случайные величины связаны линейной зависимостью, то $|K(X, Y)| = 1$.

Доказательство. Свойства 1) и 2) с очевидностью вытекают из соответствующих свойств ковариации случайных величин.

Доказательство свойства 3). Пусть $Y = AX + B$.

Имеем:

$$\begin{aligned} \text{cov}(X, Y) &= E[(X - EX)(Y - EY)] = \\ &= E[(X - EX)(AX + B - E[AX + B])] = \\ &= E[(X - EX)(AX - E[AX])] = AE[(X - EX)^2] = A DX. \end{aligned}$$

Учитывая, что (см. свойства дисперсии)

$$D[AX + B] = D[AX] + DB = A^2 DX,$$

получим утверждение свойства 3).

§ 10. Предельные теоремы теории вероятностей

Выделяются две группы предельных теорем.

¹²От латинского correlatio - взаимосвязь, относительный показатель линейной взаимосвязи между случайными переменными.

Первая группа предельных теорем получена на основе неравенства Чебышева и включает в себя теоремы Чебышева, Маркова, Бернулли, Пуассона. Эта группа теорем обычно называется *законом больших чисел*. Практический смысл этого закона: *средний результат массовых случайных явлений практически перестает быть случайным, весьма устойчив, является закономерным и его можно прогнозировать*.

Вторая группа предельных теорем (рассматривается в схеме испытаний Бернулли) состоит из теоремы Лапласа и теоремы Муавра-Лапласа и объединяется под общим названием *центральная предельная теорема*. Практический смысл этой теоремы в том, что *закон распределения суммы достаточно большого числа случайных величин приближается к нормальному закону распределения*.

Предельные теоремы имеют важное значение в теории вероятностей, являются по существу связующим звеном между теорией вероятностей и статистикой как теорией, изучающей закономерности случайных явлений при массовом наблюдении над ними.

Ниже будут рассмотрены совокупность предложений, которые с вероятностью, близкой к единице, утверждают, что наступление некоторого события зависит от совокупности неограниченного числа случайных событий (факторов), каждое из которых оказывает на него незначительное влияние, то есть результат почти не зависит от случая. Этот общий принцип носит название — закон больших чисел¹³.

Перейдем к формулировке и доказательству некоторых теорем закона больших чисел, используя метод Чебышева, основанный на точном установлении общих свойств математических

¹³Исторически первое такое предложение было сформулировано Я. Бернулли и опубликовано после его смерти в 1713 году. В 1837 году С. Пуассон доказал аналогичную теорему в более широких условиях. В 1866 году русский математик П.Л. Чебышев в работе "О средних величинах" изменил концепцию Я. Бернулли, перейдя от рассмотрения случайных событий к независимым случайным величинам. Позже А.А. Марковым были получены более общие результаты, в том числе применимость закона больших чисел к суммам зависимых величин.

ожиданий и на использовании оценки распределений случайных величин.

10.1. Неравенство Чебышева

Предположим, что задано вероятностное пространство (Ω, \mathcal{U}, P) , на котором определена случайная величина X .

Теорема 10.1.1 (Неравенство Чебышева). *Пусть для случайной величины X существует математическое ожидание EX и дисперсия DX , тогда для любого $\varepsilon > 0$*

$$P\{|X - EX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}. \quad (10.1.1)$$

Доказательство проведем для непрерывной случайной величины, заданной плотностью распределения $p(x)$. В этом случае вероятность попасть в любой интервал (a, b) определяется интегралом от плотности по отрезку $[a, b]$. Поэтому, обозначая $m = EX$, имеем:

$$P\{|X - m| \geq \varepsilon\} = \int_{\{x: |x-m| \geq \varepsilon\}} p(x) dx. \quad (10.1.2)$$

С другой стороны,

$$DX = \int_{-\infty}^{+\infty} |x - m|^2 p(x) dx \geq \int_{\{x: |x-m| \geq \varepsilon\}} |x - m|^2 p(x) dx.$$

Это неравенство только усилится, если величину $|x - m|^2$ заменить на ε^2 :

$$DX \geq \varepsilon^2 \int_{\{x: |x-m| \geq \varepsilon\}} p(x) dx.$$

Воспользуемся равенством (10.1.2), тогда

$$DX \geq \varepsilon^2 P\{|X - m| \geq \varepsilon\},$$

откуда следует (10.1.1). Доказательство закончено.

Следствие. Неравенство Чебышева, записанное в виде

$$P\{|X - m| < \varepsilon\} \geq 1 - \frac{DX}{\varepsilon^2}, \quad (10.1.3)$$

дает нижнюю границу отклонения любой случайной величины от своего математического ожидания. Это неравенство подтверждает правильность выбора дисперсии в качестве меры рассеивания значений случайной величины вокруг своего математического ожидания. В самом деле, с уменьшением дисперсии необходимо уменьшать ε , чтобы нижняя граница вероятности ($= 1 - \frac{DX}{\varepsilon^2}$) отклонения случайной величины от своего математического ожидания не изменилась.

Теперь вернемся к *правилу трех сигм*. Положим в (8.2.5) $\varepsilon = k\sigma$ при $k = 1, 2, 3$. Получим, что вероятности того, что случайная величина выйдет за пределы $k\sigma$ от математического ожидания имеет следующие верхние границы:

$$\begin{aligned} \text{при } k = 1 \quad & P\{|X - m| > \sigma\} \leq 1; \\ \text{при } k = 2 \quad & P\{|X - m| > 2\sigma\} \leq 1/4; \\ \text{при } k = 3 \quad & P\{|X - m| > 3\sigma\} \leq 1/9. \end{aligned}$$

Из этих соотношений мы видим, что правило 3σ , полученное ранее только для нормального распределения, распространяется для произвольных случайных величин.

Правило 3σ (общий случай). Вероятность того, что отклонение случайной величины от своего математического ожидания будет по абсолютной величине больше чем 3σ , не может быть больше $1/9$, т.е. практически мала.

Правило 3σ применяется в общем случае, когда закон распределения случайной величины неизвестен, а известны только m и σ_x , при этом отрезок с границами $m \pm 3\sigma_x$ считается отрезком практически возможных значений случайной величины.

В большинстве случаев, эта вероятность значительно меньше $1/9$. Так для нормального закона она равна 0,003 (см. правило 3σ для нормального распределения).

Отметим, что *неравенство Чебышева имеет практическое значение только при относительно больших ε* .

Действительно, пусть $\varepsilon = \frac{1}{2}\sqrt{D_x}$, тогда

$$P \left\{ |X - m| \geq \frac{1}{2}\sqrt{D_x} \right\} \leq \frac{D_x}{D_x/4} = 4.$$

Но ясно, что вероятность любого события не может быть больше 1. Если же $\varepsilon = 10\sqrt{D_x}$, то $P \{ |X - m| \geq 10\sqrt{D_x} \} = \frac{D_x}{100D_x} = 0,001$. Это уже неплохая оценка вероятности.

10.2. Теорема Чебышева

Все изучаемые далее случайные величины рассматриваются на одном и том же вероятностном пространстве (Ω, \mathcal{U}, P) .

Теорема 10.2.1. Пусть X_1, X_2, \dots, X_n — попарно независимые случайные величины и

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n DX_i = 0. \quad (10.2.1)$$

Тогда для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right| < \varepsilon \right\} = 1. \quad (10.2.2)$$

Доказательство. Согласно неравенству Чебышева (10.1.1), имеем:

$$P \left\{ \left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right| \geq \varepsilon \right\} \leq \frac{1}{\varepsilon^2} D \left[\frac{\sum_{i=1}^n X_i}{n} \right].$$

Так как случайные величины X_1, X_2, \dots, X_n попарно независимы, то согласно свойствам дисперсии, получим:

$$P \left\{ \left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right| \geq \varepsilon \right\} \leq \frac{1}{\varepsilon^2 n^2} \sum_{i=1}^n DX_i \rightarrow 0, \quad n \rightarrow \infty.$$

Но тогда вероятность противоположного события стремится к единице, т.е. справедливо равенство (10.2.2). Доказательство закончено.

Теорема (10.2.1) справедлива и для зависимых случайных величин. Приведем формулировку этой теоремы без доказательства.

Теорема 10.2.2 (Теорема Маркова). *Пусть последовательность X_1, X_2, \dots, X_n состоит из зависимых случайных величин и*

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n DX_i = 0.$$

Тогда для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right| < \varepsilon \right\} = 1.$$

Теорема 10.2.3 (Обобщенная теорема Чебышева). *Если X_1, X_2, \dots, X_n — попарно независимые случайные величины с математическими ожиданиями EX_1, EX_2, \dots, EX_n и $\forall i \quad DX_i < C$, то для любого $\varepsilon > 0$*

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right| < \varepsilon \right) = 1. \quad (10.2.3)$$

Доказательство. Согласно неравенству (10.1.1)

$$\begin{aligned} P \left\{ \left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right| \geq \varepsilon \right\} &\leq \\ &\leq \frac{1}{\varepsilon^2 n^2} \sum_{i=1}^n DX_i \leq \frac{1}{\varepsilon^2 n^2} n C \rightarrow 0 \end{aligned}$$

при $n \rightarrow \infty$. Но тогда вероятность противоположного события стремится к единице, т.е. справедливо равенство (10.2.3). Доказательство закончено.

Обобщенная теорема Чебышева является одной из важнейших в группе теорем — закона больших чисел. Она устанавливает связь между средним арифметическим наблюдаемых значений случайной величины и ее математическим ожиданием.

Пределы в теоремах (10.2.2) и (10.2.3) удобно формулировать, используя понятие *сходимости по вероятности*.

Определение 10.2.1. Последовательность случайных величин X_1, X_2, \dots, X_n сходится по вероятности к величине a , $(X_n \xrightarrow{P} a)$, если для любого $\varepsilon > 0$ $\lim_{n \rightarrow \infty} P\{|X_n - a| < \varepsilon\} = 1$.

10.3. Закон больших чисел

Здесь мы рассмотрим одну из простейших и вместе с тем наиболее важных форм закона больших чисел.

Предварительно проведем следующие вспомогательные рассуждения.

Имеется случайная величина X с математическим ожиданием $EX = m$ и с дисперсией $DX = d$. Производится n опытов, в каждом из которых эта случайная величина может принять одно из своих значений. Через X_i обозначим случайную величину — реализация случайной величины X в i -ом опыте. Предположим, что произведено n опытов, и рассмотрим новую случайную величину Y — среднее арифметическое этих величин:

$$Y = \frac{\sum_{i=1}^n X_i}{n}.$$

Ясно, что все случайные величины X_i имеют те же числовые характеристики, что и случайная величина X . Найдем математическое ожидание и дисперсию для Y . Имеем:

$$EY = \frac{1}{n} \sum_{i=1}^n EX_i = \frac{1}{n} \sum_{i=1}^n m = \frac{1}{n} nm = m,$$

$$DY = \frac{1}{n^2} \sum_{i=1}^n DX_i = \frac{1}{n^2} \sum_{i=1}^n d = \frac{1}{n^2} nd = \frac{d}{n}.$$

Из этих вспомогательных рассуждений получим следующие выводы:

— математическое ожидание "средней" случайной величины Y не зависит от числа опытов и равно математическому ожиданию наблюдаемой случайной величины X .

— дисперсия "средней" случайной величины Y неограниченно убывает с увеличением числа опытов (чем больше опытов,

тем меньше отклонение (разброс) значений случайной величины X от своего математического ожидания).

Последнее требование как раз и содержится в теореме Чебышева. Теперь сформулируем **закон больших чисел**.

При достаточно большом числе независимых опытов среднее арифметическое наблюдаемых значений случайной величины сходится по вероятности к ее математическому ожиданию:

$$\lim_{n \rightarrow \infty} P \left\{ \left| \sum_{i=1}^n \frac{X_i}{n} - m \right| < \varepsilon \right\} = 1.$$

Или среднее арифметическое реализаций случайных величин сходится по вероятности к ее математическому ожиданию при неограниченном увеличении числа опытов, т.е.

$$\left(\frac{\sum_{j=1}^n X_j}{n} \right) \xrightarrow{p} m.$$

Точная формулировка закона больших чисел:

Теорема 10.3.1 (Теорема Хинчина). *Если X_1, X_2, \dots, X_n — одинаково распределенные и попарно независимые случайные величины, имеют конечную дисперсию $DX_i = \sigma^2 < \infty$ и математическое ожидание $EX_i = m$, то для любого $\varepsilon > 0$*

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\sum_{i=1}^n X_i}{n} - m \right| < \varepsilon \right\} = 1. \quad (10.3.1)$$

Доказательство с очевидностью вытекает из теоремы Чебышева,

$$\text{т.к. } \frac{\sum_{i=1}^n EX_i}{n} = \frac{na}{n} = m.$$

Теорема 10.3.2 (Теорема Бернулли). *Пусть μ_n — число появлений события A в схеме испытаний Бернулли и $P(A) = p$, тогда для любого $\varepsilon > 0$*

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1. \quad (10.3.2)$$

Доказательство. Обозначим через X_i случайную величину, равную

$$\begin{cases} 1, & \text{если в } i\text{-ом испытании событие } A \text{ произошло} \\ 0, & \text{если в } i\text{-ом испытании событие } A \text{ не произошло.} \end{cases}$$

Тогда $\mu_n = \sum_{i=1}^n X_i$. Кроме того,

$$m = a = E[X_i] = 1p + 0(1-p) = p, \quad (\forall i = 1, 2, 3, \dots).$$

Отсюда

$$P\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} = P\left(\left|\frac{\sum_{i=1}^n X_i}{n} - a\right| < \varepsilon\right).$$

Теперь утверждение (10.3.2) следует из равенства (10.3.1).

Доказательство закончено.

Как было отмечено выше, теорема Бернулли является исторически первой и важнейшей формой закона больших чисел. Она устанавливает связь между частотой события $\frac{\mu_n}{n}$ и ее вероятностью p : относительная частота $\frac{\mu_n}{n}$ появления события A в каждом испытании стремится к вероятности этого события при неограниченном увеличении количества испытаний. Рассмотрим далее обобщение теоремы Бернулли — теорему Пуассона, допускающую в отличие от теоремы Бернулли зависимость вероятности появления события $P(A) = p_i$ от номера испытания. Эту теорему приведем в качестве упражнения.

♦ **Пример.**

Теорема 10.3.3 (Теорема Пуассона). *При неограниченном увеличении числа независимых испытаний в переменных условиях частота события A сходится по вероятности к среднему арифметическому вероятностей p_i появления события A в i -ом испытании:*

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{\mu_n}{n} - \frac{\sum_{i=1}^n p_i}{n}\right| < \varepsilon\right) = 1.$$

Указание. Используйте случайную величину X_i , введенную при доказательстве теоремы Пуассона, и обобщенную теорему Чебышева.

10.4. Усиленный закон больших чисел

Пусть X_1, X_2, \dots, X_n — последовательность случайных величин с конечной дисперсией. Если дисперсии равномерно ограничены, то при $n \rightarrow \infty$ имеет место закон больших чисел

$$\left(\frac{\sum_{i=1}^n X_i}{n} \right) \xrightarrow{p} a$$

или говорят, что последовательность случайных величин $\{X_n, n \geq 1\}$ удовлетворяет закону больших чисел.

Прежде чем дать формулировку усиленного закона больших чисел, определим еще один вид сходимости — сходимость с вероятностью единица (или сходимость почти наверное).

Определение 10.4.1. Последовательность случайных величин X_1, X_2, \dots, X_n сходится почти наверное к случайной величине X , $(X_n \xrightarrow{n.н.} X)$, если

$$P\left\{\lim_{n \rightarrow \infty} X_n = X\right\} = 1.$$

Усиленным законом больших чисел¹⁴ называют утверждения, в которых сходимость по вероятности заменяется сходимостью с вероятностью единица. Таким образом, если в усиленном законе больших чисел говорят о поведении всей последовательности сумм в целом, то в обычном законе больших чисел речь идет лишь об отдельных суммах. При этом отдельно рассматривают случаи, когда случайные величины $\{X_n, n \geq 1\}$ распределены одинаково и имеют произвольное распределение.

Сформулируем некоторые теоремы усиленного закона больших чисел для последовательности случайных величин.

¹⁴Впервые был сформулирован и доказан в 1909 году французским математиком Э. Борелем (1871–1956) для схемы Бернулли в случае $p = \frac{1}{2}$. Ф. Кантели установил достаточные условия усиленного закона больших чисел для независимых случайных величин в терминах вторых и четвертых центральных моментов слагаемых. Дальнейшее расширение условий применимости усиленного закона больших чисел было осуществлено А.Я. Хинчиным (ввел термин "усиленный закон больших чисел"), А.Н. Колмогоровым и Ю.В. Прохоровым.

Теорема 10.4.1 (Теорема Кантелли). Пусть X_1, X_2, \dots, X_n — последовательность независимых случайных величин с конечным четвертым моментом и такие, что для некоторого $C = \text{const}$

$$E|X_n - EX_n|^4 \leq C, \quad n \geq 1.$$

Тогда при $n \rightarrow \infty$

$$\left(\frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right) \xrightarrow{n.н.} 0. \quad (10.4.1)$$

Теорема 10.4.2 (Теорема Колмогорова). Пусть X_1, X_2, \dots, X_n — последовательность независимых случайных величин с конечным вторым моментом, т.е. $DX_n = \sigma^2 < \infty$ и выполнено условие (Колмогорова)

$$\sum_n \frac{DX_n}{n^2} < \infty.$$

Тогда при $n \rightarrow \infty$

$$\left(\frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n EX_i}{n} \right) \xrightarrow{n.н.} 0. \quad (10.4.2)$$

Замечание 10.4.1. Условие Колмогорова $\sum_n \frac{DX_n}{n^2} < \infty$ является достаточным, но не необходимым для справедливости усиленного закона больших чисел.

В теории известно много примеров последовательностей независимых случайных величин, для которых имеет место усиленный закон больших чисел, но условие Колмогорова не выполнено (см. [17], стр. 154.).

В случае, если случайные величины взаимно независимы и одинаково распределены, для подчинения их усиленному закону больших чисел нет необходимости существования второго момента. Необходимым и достаточным условием для применимости усиленного закона больших чисел для такой последовательности случайных величин является существование математического ожидания.

Теорема 10.4.3 (Теорема Колмогорова). Пусть X_1, X_2, \dots, X_n — последовательность независимых, одинаково распределенных случайных величин с $E|X_n| < \infty$. Тогда при $n \rightarrow \infty$

$$\left(\frac{\sum_{i=1}^n X_i}{n} \right) \xrightarrow{\text{п.н.}} m, \quad (10.4.3)$$

где $m = EX_n$.

Замечание 10.4.2. В схеме Бернулли для числа успехов имеет место не только закон больших чисел $\frac{\mu_n}{n} \xrightarrow{p} p$, но и усиленный закон больших чисел $\frac{\mu_n}{n} \xrightarrow{\text{п.н.}} p$.

Замечание 10.4.3. Утверждение теоремы можно сформулировать следующим образом. Пусть X_1, X_2, \dots, X_n — последовательность независимых, одинаково распределенных случайных величин, для которых

$$\left(\frac{\sum_{i=1}^n X_i}{n} \right) \xrightarrow{\text{п.н.}} C,$$

где C — конечная константа. Тогда $E|X_n| < \infty$ и $C = EX_n$.

Замечание 10.4.4. Если математическое ожидание существует, но не обязательно конечно, то утверждение теоремы (10.4.3) остается в силе.

Доказательства сформулированных выше теорем читатель может найти, например, в следующих книгах: [4], [16].

Закон больших чисел и усиленный закон больших чисел имеют большое не только теоретическое и методологическое значение для науки, но и огромную, прежде всего, практическую ценность. Они используются при математическом изучении явлений природы, физических, химических, технических процессов, а также тех или иных общественных явлений. Проиллюстрируем действие законов следующим схематическим примером.

♦ **Пример 1.** Рассмотрим отрезок проводника (пусть это будет кусок медной проволоки) определенного сечения. Заранее нельзя сказать, с какой скоростью и в каком направлении будут двигаться отдельные свободные электроны, так как их движение хаотично. Но если рассмотреть поведение электронов при

определенных условиях — наложение внешнего электрического поля на проводник, то можно рассчитать долю свободных электронов, которые будут двигаться в определенном направлении. Именно поведение совокупности действия свободных электронов, а не каждого в отдельности определяет основное свойство проводника — проводимость электрического тока. Возможность проводимости тока определяется количеством свободных электронов, проходящих за единицу времени через единицу площади сечения проводника. Количество электронов, проходящих через сечение, меняется в зависимости от случая, однако в силу действия закона больших чисел — сила тока остается постоянной величиной. Это "уравнивающее" влияние закона больших чисел выполняется с высокой точностью. Отсутствие тока в проводнике в обычных условиях (без наложения электрического поля) можно объяснить с позиций усиленного закона больших чисел. Определенная доля свободных электронов проходит через заданное сечение в определенный момент времени в одном направлении, другая доля — в противоположном. Из курса физики известно, что оба эти направления компенсируются. Действие различных случайных факторов вызывает небольшое отклонение в ту или иную сторону, однако суммарное поведение свободных электронов становится вполне закономерным, тока в проводнике не может быть.

Глава 3

Элементы теории случайных процессов

Следуя исторической концепции, с тем чтобы объяснить возникновение и развитие основных положений теории вероятностей, дадим краткое обоснование появлению новой теории — теории случайных процессов. В результате дальнейшего изучения реальных процессов, но теперь уже во времени, многие ученые, в том числе физики, биологи, инженеры и другие, подошли к весьма важной проблеме. Теория вероятностей предлагала им в качестве математического аппарата лишь средства, изучавшие стационарные состояния. Необходима же была теория, которая изучала случайные величины, зависящие от одного или нескольких непрерывно изменяющихся параметров. В XX веке проблема достигла своей зрелости, вызвав эвристическое направление в поисках ее решения.

Начало общей теории случайных процессов было положено работами советских математиков А. Н. Колмогоровым (1903–1987), А. Я. Хинчиным (1894–1959), Е. Е. Слуцким (1880–1948), Н. Винером (1894–1965). Однако у них были предшественники — П. Лаплас, Л. Башелье, Ж. Пуанкаре, А.А. Мар-

ков. Колмогоровым было дано систематическое и строгое построение основ теории стохастических процессов без последствий (процессов марковского типа). В работах А.Я. Хинчина была создана теория стационарных процессов. Н. Винер в середине двадцатых годов при изучении броуновского движения ввел в рассмотрение процесс, получивший название винеровского процесса. Работы Е. Е. Слуцкого (1880–1948) посвящены изучению теории случайных функций. В наши дни теория случайных процессов занимает центральное место не только в теории вероятностей, но также в естествознании, в инженерном деле, в экономике, в организации производства, в теории связи, в теории массового обслуживания и др..

Прежде чем перейти к изучению элементов теории случайных процессов, рассмотрим некоторые задачи, которые явились исходным пунктом новой теории.

1. Изучение процесса диффузии

Рассмотрим поведение какой-нибудь молекулы газа или жидкости. Эта молекула в случайные моменты сталкивается с другими молекулами и меняет при этом направление движения и скорость. Состояние молекулы, таким образом, подвержено случайным изменениям и представляет собой ничто иное, как случайный процесс, который определяется шестью параметрами - тремя координатами и тремя компонентами скорости. Как быстро протекает процесс диффузии, по каким законам, когда образующая смесь становится однородной? На все эти вопросы дает ответ статистическая теория диффузии¹, в основе которой лежит теория случайных процессов. Подобные задачи возникают в химии при изучении химических реакций. Теория случайных процессов дает ответы на вопросы: какая часть молекул уже вступила в реакцию, какова особенность протекания реакции во времени и др.?

¹Попытка изучения средствами теории вероятностей явления диффузии была предпринята в 1914 г. двумя известными физиками — М. Планком (1858–1947) и А. Фоккером (1887–1972).

2. Изучение явлений, протекающих по принципу радиоактивного распада

Многочисленные наблюдения показывают, что распад отдельных атомов происходит в случайно взятые моменты времени и расположение этих моментов, если количество распадающегося вещества не превосходит некоторого определенного критического предела, не зависит друг от друга. Для изучения процесса радиоактивного распада весьма важно определить вероятность того, что за определенный промежуток времени распадается то или иное число атомов. Аналогично происходят многие другие процессы: процесс броуновского движения, который определяет число броуновских частиц, оказавшихся в данный момент в определенной области пространства. Изучением таких процессов занимается статистическая теория броуновского движения². Или различные процессы массового обслуживания, например, число вызовов от абонентов, поступающих на телефонную станцию, количество обращений в банкоматы, терминалы оплаты сотовой связи и др., базовой основой которых является также теория случайных процессов.

3. Исследования, связанные с фондовой биржей

Одним из способов ограничения массы денег, обращающихся в стране, является использование различных векселей, расписок и других ценных бумаг, которые монопольно выпускаются государством. Известно, что акции ликвидны, то есть могут быть проданы на рынке — превращены в ту или иную валюту. Как правильно выбрать стратегию биржевой игры на основе прошлых данных о ценах закрытия, как проследить динамику цен акций за какой-то период, как удачно выбрать момент покуп-

²Теория броуновского движения, исходящая из теоретико-вероятностных предпосылок, была разработана в 1905 г. двумя известными физиками М. Смолуховским (1872–1917) и А. Эйнштейном (1879–1955). Основные положения их трудов использовались неоднократно как при изучении физических явлений, так и в различных инженерных задачах. В частности, именно с этих работ, а также с работ Эрланга, проявился широкий интерес к процессу Пуассона. Впрочем, сам Пуассон ввел в рассмотрение только распределение Пуассона, но он заслужил, чтобы его имя произносилось и при изучении случайных процессов, связанных с его распределением.

ки и продажи акций? На эти вопросы также дает ответ теория случайных процессов.

§ 11. Введение в теорию случайных процессов

11.1. Определение случайного процесса

В курсе теории вероятностей основной аксиоматики является понятие вероятностного пространства (Ω, \mathcal{U}, P) , где Ω — пространство элементарных событий, \mathcal{U} — σ -алгебра подмножеств Ω , называемая событиями, P — вероятностная мера на \mathcal{U} . Случайной величиной X , принимающей значения в множестве B , называется любая функция $X : \Omega \rightarrow B$, измеримая относительно \mathcal{U} . Обычно за множество B принимают $B = Z$, или $B = R_1$, или $B = R_n$ (в случае векторных случайных величин или геометрических вероятностей). В теории случайных процессов используется та же конструкция и практически та же терминология, определенная для случайных величин³. Введем понятие случайного процесса, базирующегося на аксиоматике теории вероятностей.

Определение 11.1.1. *Случайным процессом $\{X(t, \omega)\}$ называют последовательность или совокупность случайных величин, заданных на вероятностном пространстве (Ω, \mathcal{U}, P) , зависящих от параметра t , $t \in T$ (t интерпретируется как время).*

³Только в качестве множества B выбирается множество более сложной структуры. Например, если $B = Z^z$ или $B = R^z$, то случайная величина X принимает значения в пространстве целочисленных или действительных последовательностей, и говорят, что X — случайный процесс с дискретным временем.

Определение 11.1.2. *Реализацией (траекторией) случайного процесса называется неслучайная функция, полученная при фиксированном $\omega_0 \in \Omega$. Совокупность всех возможных реализаций называют ансамблем реализаций данного случайного процесса.*

Определение 11.1.3. *Сечением случайного процесса называется случайная величина $X(t_0, \omega)$, полученная при фиксированном $t_0 \in T$.*

Таким образом, случайный процесс можно рассматривать либо как совокупность случайных величин $X(t)$, зависящих от параметра t , либо как совокупность реализаций процесса $X(t)$. При этом для определения случайного процесса необходимо задавать вероятностную меру в пространстве реализаций процесса.

11.2. Классификация случайных процессов

Основные признаки, по которым различаются случайные процессы, касаются природы пространства состояний (обозначаемого через E), временного параметра T и отношений зависимости между случайными величинами.

Определение 11.2.1. *Пространство состояний — это пространство, которому принадлежат все возможные значения, принимаемые всеми случайными величинами $X(t)$ или X_t . Если E совпадает с R , то процесс $\{X(t)\}$ называется действительным случайным процессом. Если E — евклидово k -мерное пространство, то процесс $\{X(t)\}$ является k -мерным.*

Определение 11.2.2. *Если $T = \{0, 1, 2, \dots\}$, то $\{X(t)\}$ является процессом с дискретным временем. Если $T = [0; \infty)$, то $\{X(t)\}$ — процесс с непрерывным временем.*

♦ **Пример 1.** Любой выборочный контроль продукции будет относиться к случайным процессам с дискретными состояниями (e_1 — годная, e_2 — негодная продукция) и дискретным временем (t_1, t_2 — время проверки). С другой стороны, случай отказа любой машины можно отнести к случайным процессам с дискретными состояниями, но непрерывным временем. Регистрацию температуры воздуха в определенные моменты времени можно отнести к

случайному процессу с непрерывным состоянием и дискретным временем, в свою очередь случайный процесс изменения напряжения в электросети питания ЭВМ является примером случайного процесса с непрерывным состоянием и временем.

Важной чертой случайного процесса $\{X(t)\}$ является зависимость между случайными величинами $X(t)$. Характер этой зависимости определяется заданием совместных функций распределений для каждого конечного семейства $X(t_1), X(t_2), \dots$. В теории случайных процессов выделяют различные широкие классы случайных процессов, которые могут и пересекаться. Опишем некоторые классические типы случайных процессов, характеризующиеся различными видами зависимости между $X(t)$, $t \in T$.

1. Марковские процессы

Являются самым крупным направлением в теории случайных процессов.

Случайный процесс $\{X(t)\}$ называется марковским, если выполняется условие Маркова: $\forall t_1 < t_2 < \dots < t_n$ справедливо соотношение

$$P\{X(t_n) = e_n / X(t_1) = e_1, X(t_2) = e_2, \dots, X(t_{n-1}) = e_{n-1}\} = \\ = P\{X(t_n) = e_n / X(t_{n-1}) = e_{n-1}\}, \quad \text{где } e_1, e_2, \dots, e_n \in E.$$

Для 3-х моментов времени $t_i < t_j < t_k$ условие Маркова примет вид:

$$P\{X(t_k) = e_k / X(t_j) = e_j, X(t_i) = e_i\} = P\{X(t_k) = e_k / X(t_j) = e_j\}.$$

Таким образом, будущее состояние системы зависит от прошлого только через настоящее. Марковские процессы являются естественным обобщением детерминированных процессов, рассматриваемых в классической физике. В детерминированных процессах состояние системы в момент времени t_0 однозначно определяет ход процесса в будущем; в марковских процессах состояние системы в момент времени t_0 однозначно определяет распределение вероятностей хода процесса при $t > t_0$, причем никакие сведения о ходе процесса до момента времени t не изменяют это распределение. Марковские процессы получили широкое

применение в статистическом контроле качества продукции, в теории массового обслуживания, в теории надежности, в теории планирования.

2. Стационарные случайные процессы

Являются вторым по значимости классом случайных процессов.

Случайный процесс $\{X(t)\}$, $t \in T \leq R_1$ называется стационарным в узком смысле, если совместные распределения семейств $X(t_1+h), X(t_2+h), \dots, X(t_n+h)$ и $X(t_1), X(t_2), \dots, X(t_n)$ одинаковы при всех $n > 0$ и всех $t_1, t_2, \dots, t_n \in T$, т.е. многомерная функция распределения не зависит от начала отсчета или от сдвига всех сечений вправо или влево на один и тот же интервал времени.

Случайный процесс $\{X(t)\}$, $t \in T \leq R_1$ называется стационарным в широком смысле, если его математическое ожидание и дисперсия не зависят от времени, а ковариационная функция зависит только от разности моментов времени, т.е. $B(t, t+h) = B(h)$, где $B(t, t+h) = cov(X(t, \omega), X(t+h, \omega))$.

Узкое и широкое определения стационарности не тождественны. Случайные процессы, стационарные в узком смысле, всегда стационарны в широком смысле, но не наоборот. Схема стационарных случайных процессов с хорошим приближением описывает многие реальные явления, сопровождающиеся неупорядоченными флуктуациями. Так, например, пульсации силы тока или напряжения в электрической цепи (электрический "шум") можно рассматривать как стационарный случайный процесс, если эта цепь находится в стационарном режиме, т.е. если все ее макроскопические характеристики и все условия, вызывающие протекание через нее тока, не меняются во времени; пульсации скорости в точке турбулентного течения представляют собой стационарный случайный процесс, если не меняются общие условия, порождающие рассматриваемое течение (т.е. течение является установившимся), и т.д.. Выделение понятия стационарного случайного процесса и получение первых относящихся к нему математических результатов, как было отмечено выше, являются заслугой Е.Е. Слуцкого и относятся к концу 20-х и началу 30-х

гг. XX века. В дальнейшем важные работы по теории стационарных случайных процессов были выполнены А.Я. Хининым, А.Н. Колмогоровым, Г. Крамером, Н. Винером и др.

3. Мартингалы

В последние десятилетия интенсивно ведутся исследования класса случайных процессов, которые называются мартингалами⁴. Общность этого класса определяется тем, что на характер зависимости между случайными величинами накладываются довольно слабые условия, выражаемые в терминах условных математических ожиданий.

Пусть $\{X(t)\}$, $t \in T$ — действительный случайный процесс с дискретным или непрерывным параметром T . Случайный процесс $\{X(t)\}$ называется мартингалом, если для $\forall t_1 < t_2 < \dots < t_{n+1}$, $t_i \in T$ справедливо $E\{X(t_{n+1})/X(t_1) = e_1, \dots, X(t_n) = e_n\} = e_n$, для всех допустимых значений e_1, e_2, \dots, e_n , то есть условное математическое ожидание мартингала в следующий момент времени равно его значению в предыдущий момент времени. Иными словами, мартингал — это процесс без сноса (в том смысле, что его математическое ожидание в следующий момент времени совпадает со значением в текущий момент).

♦ **Пример 2.** Если $X(t)$ описывает состояние капитала игрока в момент t , то по определению мартингала средняя величина его капитала в момент t_{n+1} при условии, что в момент t_n он располагал капиталом e_n , равно e_n , независимо от того, каков был его капитал в предшествующие моменты времени.

В математической теории мартингалов основную роль играют моменты остановки ν . В математических моделях финансового рынка процесс $\{X(t)\}$ может описывать, например, цену акций в момент t , а случайный момент ν соответствовать моменту продажи пакета его владельцем. Мартингалы с успехом применяются в генетике, теории потенциала, стохастических интегралах и т.д.

⁴Французский финансовый математик Л. Башелье создал общую математическую теорию безобидных игр — мартингал, которая позднее, после исследований Ж. Вилле, П. Леей, Д. Дуба (1910–2004) и др., стала одним из важнейших классов стохастических процессов.

4. Процесс с независимыми приращениями

Именно с изучения данного класса возникла теория случайных процессов. Вначале изучались винеровские процессы (процессы броуновского движения⁵), затем более общие процессы с независимыми приращениями.

Случайный процесс $\{X(t)\}$, $t_i \in T \leq R_1$ — процесс с независимыми приращениями, если его приращения на неперекрывающихся отрезках не зависят друг от друга: для $\forall t_0 < t_1 < \dots < t_n \in T$ приращения случайных величин $X(t_1) - X(t_0)$, $X(t_2) - X(t_1)$, \dots , $X(t_n) - X(t_{n-1})$, независимы.

Примерами процессов с независимыми приращениями являются процесс броуновского движения (процесс с независимыми приращениями, для которых распределение величины $X(t+h) - X(t)$, $t \geq 0, h > 0$ является гауссовым), процесс Пуассона (стохастически непрерывный процесс с независимыми приращениями, если $X(t) - X(s)$, $\forall t, s \geq 0, s < t$), безгранично делимые распределения и др.

§ 12. Математический аппарат дискретных марковских цепей

Цепь Маркова — одна из первых строго математически обоснованных моделей случайных процессов. Такой процесс был

⁵Представляет собой модель движения маленькой частицы, взвешенной в жидкости или газе и совершающей хаотическое движение в результате столкновения с молекулами окружающей среды. Впервые такое явление наблюдал голландский натуралист (известен как изобретатель микроскопа) Антони ван Левенгук (1632–1723) в 1670 году, изучая поведение частицы пыльца в водном растворе под микроскопом. Английский ботаник Р. Броун (1773–1858) в 1828 г. подробно описал свойства этого движения. Математически строгую теорию броуновского движения построил в 1923 г. Н. Винер (1894–1964), однако значение его работы было понято после создания аксиоматической теории вероятностей А.Н. Колмогоровым.

впервые рассмотрен в работе А.А. Маркова (1856–1922), посвященной анализу последовательности гласных и согласных букв в русском тексте. Общая теория марковских процессов и их классификация была дана А.Н. Колмогоровым в 1930 году. Его исследования дали логически безупречную математическую основу общей теории марковских процессов. В настоящее время теория марковских процессов и ее приложения широко применяются в самых различных областях наук: в механике, физике, химии и др. Благодаря сравнительной простоте и наглядности математического аппарата, высокой достоверности и точности получаемых решений, особое внимание марковские процессы приобрели у специалистов, занимающихся исследованием операций и теорией принятия оптимальных решений.

12.1. Классификация марковских случайных процессов

Марковские процессы различаются состояниями во времени и в пространстве. В отношении состояния процесса во времени различают два типа марковских процессов: с дискретным временем и непрерывным временем.

Определение 12.1.1. *Марковским случайным процессом с дискретным временем называется такой процесс, у которого переходы из одного состояния в другое возможны в строго определённые, заранее заданные моменты времени t_1, t_2, \dots, t_k , называемые шагами процесса. Такой процесс называется цепью Маркова или процессом с дискретным параметром T , который может быть процессом с конечным или бесконечным множеством состояний.*

Определение 12.1.2. *Марковским случайным процессом с непрерывным временем называется такой процесс, у которого переход из одного состояния в другое возможен в любой момент времени t .*

Случайные процессы, в том числе марковские, могут быть с дискретным числом состояний. В этом случае они называются **дискретными случайными процессами**. Число состоя-

ний является счётным конечным или бесконечным. И выделяют класс **непрерывных случайных процессов**, для которых возможно бесконечное множество состояний.

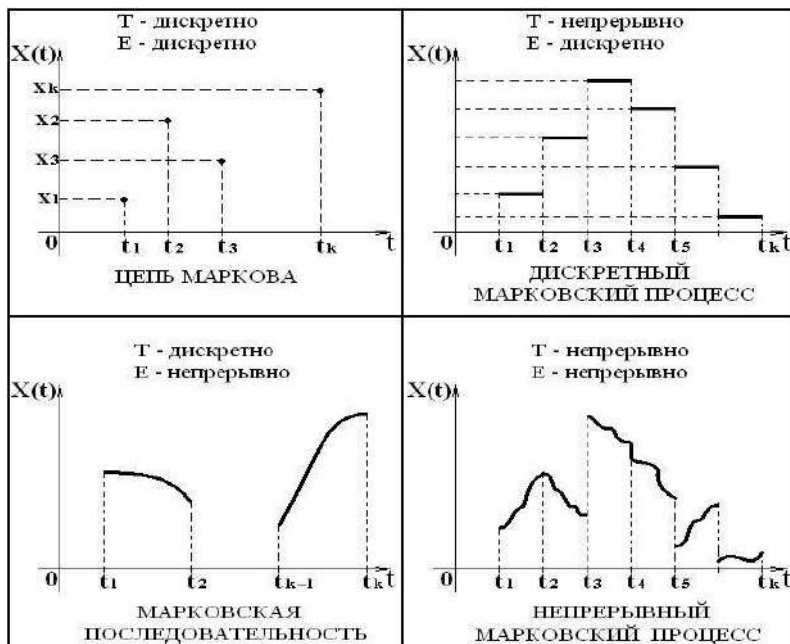


Рис. 20

12.2. Вероятностные характеристики цепей Маркова

Пусть $E = \{e_0, e_1, \dots, e_k, \dots\}$ — некоторое конечное или счетное множество. Рассмотрим последовательность случайных величин $\{X_n\}$, $n = 0, 1, \dots$, которые принимают значения из заданного множества с вероятностями $P_k(n) = P\{X_n = e_k\}$, $k = 0, 1, \dots$. Таким образом, X_n — случайная величина с дискретным (конечным или счетным) множеством состояний E . Случайную последовательность $\{X_n\}$, $n = 0, 1, \dots$ указанного типа называют **дискретной цепью**.

Определение 12.2.1. Случайная последовательность $\{X_n\}$ называется дискретной цепью Маркова ⁶, если она является дискретной цепью и обладает марковским свойством, т.е. для любых $n \geq 1$ и любых элементов e_0, e_1, \dots, e_n множества E выполнено $P\{X_n = e_n / X_{n-1} = e_{n-1}, \dots, X_0 = e_0\} = P\{X_n = e_n / X_{n-1} = e_{n-1}\}$.

Из определения следует, что дискретная цепь Маркова является частным случаем марковского процесса. Будем считать, что пространство E — измеримое множество, в котором все одноточечные множества также измеримы. Если в момент $n \geq 0$ произошло событие $\{X_n = e_k\}$, то говорят, что цепь Маркова находится в состоянии e_k . Если же известно, что для всех $n \geq 1$ выполнено $\{X_{n-1} = e_k\}$ и $\{X_n = e_{k+1}\}$, то говорят, что цепь на n -м шаге перешла из состояния e_k в состояние e_{k+1} . Если E имеет конечное число состояний, то соответствующая цепь Маркова называется **конечной**.

Определение 12.2.2. Вероятность случайной величины X_n попасть в состояние j , если известно, что X_{n-1} находится в состоянии i , называется одношаговой переходной вероятностью и обозначается p_{ij} .

По определению: $p_{ij} = P\{X_n = j / X_{n-1} = i\}$, $i, j \in E$.

Определение 12.2.3. Матрица \mathcal{P} , элементами которой являются вероятности перехода p_{ij} , называется переходной матрицей цепи Маркова $\{X_n\}$ за один шаг.

$$\mathcal{P} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}.$$

⁶Для последовательности испытаний в 5.1. было введено определение цепи Маркова. В данном разделе меняется только терминология для изучения цепей Маркова в различных приложениях. Выражение "результатом n -го испытания оказалось событие E_n " заменяется на следующее — "в момент времени n система находится в состоянии E_n "; термин "условная вероятность p_{ij} " заменяется на термин "вероятность перехода системы из состояния E_i в состояние E_j ".

Переходная матрица обладает свойствами:

- 1) $0 \leq p_{ij} \leq 1$, для любых $i, j = 0, 1, \dots$
- 2) $\sum_{j=0}^{\infty} p_{ij} = 1$, $\forall i = 0, 1, \dots$

Матрица, удовлетворяющая свойствам 1,2, называется **стохастической**. Стохастические матрицы, удовлетворяющие дополнительно условию — суммы элементов в каждом столбце равны 1, называются **дважды стохастическими**.

Определение 12.2.4. *Цепь Маркова называется однородной, если для всех $n \geq 1$ переходные вероятности не зависят от номера испытания (шага), т.е. остаются постоянными в ходе процесса: $p_{ij}(n) = p_{ij}$.*

Определение 12.2.5. *Вероятность $p_k(n) = P\{X_n = e_k\}$, $e_k \in E$ называется вероятностью состояния e_k в момент времени $n \geq 0$. Вектор $p(n) = (p_0(n), p_1(n), \dots)$ называется распределением вероятностей состояний цепи Маркова $\{X_n\}$ в момент $n \geq 0$.*

Очевидно, что $p(n)$ удовлетворяет при каждом $n < \infty$ условию нормировки: $\sum_{k=0}^{\infty} p_k(n) = 1$. Компоненты вектора $p(n)$ показывают, какие из возможных состояний цепи Маркова в момент n являются наиболее вероятными, а какие нет. Процесс полностью определен, если задана матрица переходных вероятностей и вероятности начальных состояний. Таким образом, пара $(P, p(0))$ полностью описывает вероятностную структуру однородной цепи Маркова.

Действительно, пусть $P\{X_0 = e_0\} = p_{i_0}$.

По определению условной вероятности имеем:

$$\begin{aligned} & P\{X_0 = e_0, X_1 = e_1, \dots, X_n = e_n\} = \\ & = P\{X_n = e_n / X_0 = e_0, X_1 = e_1, \dots, X_{n-1} = e_{n-1}\} \times \\ & \times P\{X_0 = e_0, X_1 = e_1, \dots, X_{n-1} = e_{n-1}\}. \end{aligned}$$

Но по определению марковского процесса:

$$\begin{aligned} & P\{X_n = e_n / X_0 = e_0, X_1 = e_1, \dots, X_{n-1} = e_{n-1}\} = \\ & = P\{X_n = e_n / X_{n-1} = e_{n-1}\} = p_{i_{n-1} i_n}. \end{aligned}$$

Используя последние два равенства, получим:

$$\begin{aligned} P\{X_0 = e_0, X_1 = e_1, \dots, X_n = e_n\} = \\ = P\{X_n = e_n / X_0 = e_0, X_1 = e_1, \dots, X_{n-1} = e_{n-1}\} \cdot p_{i_{n-1}i_n}. \end{aligned}$$

Используя индукцию, окончательно имеем:

$$P\{X_0 = e_0, X_1 = e_1, \dots, X_n = e_n\} = p_{i_0} \prod_{k=1}^n p_{i_{k-1}i_k} \quad (12.2.1)$$

или в векторной форме: $p(n) = p(0)\mathcal{P}^n$.

Замечание 12.2.1. Если $p_{kj}^{(m)}(n) = P\{X_{n+m} = e_j / X_n = e_k\}$ — вероятность перехода из состояния e_k в состояние e_j за $m \geq 1$ шагов, то для однородной цепи Маркова имеем:

$$p(n+m) = p(n)\mathcal{P}^m.$$

Первая задача в теории цепей Маркова состоит в определении вероятности перехода из состояния i в состояние j за m шагов. Пусть $\{X_n\}$ — цепь Маркова с матрицей переходных вероятностей \mathcal{P} и множеством состояний E . Матрица переходных вероятностей определяет вероятности перехода за 1 шаг. Пользуясь формулой (11.2.1), можно найти вероятности переходов за m шагов $P\{X_{n+m} = j / X_n = i\} = p_{ij}^{(m)}$, суммируя вероятности по всем возможным значениям.

Теорема 12.2.1. Для однородной цепи Маркова с матрицей переходных вероятностей \mathcal{P} при любом $m \geq 1$ справедливо равенство: $\mathcal{P}^{(m)} = \mathcal{P}^m$.

Доказательство. Так как по формуле полной вероятности при $\forall k \in \{1, 2, \dots, m-1\}$ и $\forall i, j \in E$

$$p_{ij}^{(m)} = \sum_{r=1}^n p_{ir}^{(k)} p_{rj}^{(m-k)} \quad (12.2.2)$$

то согласно (12.2.2) между матрицей с различными индексами существует соотношение $\mathcal{P}^{(m)} = \mathcal{P}^{(k)}\mathcal{P}^{(m-k)}$.

В частности, при $m = 2$ находим, что

$$\mathcal{P}^{(2)} = \mathcal{P}^{(1)}\mathcal{P}^{(1)} = (\mathcal{P}^{(1)})^2 = \mathcal{P}^2,$$

при $m = 3$:

$$\mathcal{P}^{(3)} = \mathcal{P}^{(1)}\mathcal{P}^{(2)} = (\mathcal{P}^{(1)})^3 = \mathcal{P}^3.$$

И при $\forall m$ получим: $\mathcal{P}^{(m)} = \mathcal{P}^m$.

Доказательство (12.2.1) закончено.

Допустим, что в системе протекает марковский дискретный процесс с дискретным временем. Пусть $1, 2, \dots, n$ — возможные состояния системы и t_1, t_2, \dots, t_k — шаги, в которые система может переходить из состояния в состояние, то есть имеем цепь Маркова.

Определение 12.2.6. *Цепь Маркова называется неоднородной, если переходные вероятности (хотя бы одна) зависят от номера испытания (шага) k .*

В этом случае переходные вероятности будем обозначать $p_{ij}(k)$. Тогда матрица переходных вероятностей будет зависеть от k :

$$\mathcal{P}(k) = \begin{pmatrix} p_{11}(k) & p_{12}(k) & \dots & p_{1n}(k) \\ p_{21}(k) & p_{22}(k) & \dots & p_{2n}(k) \\ \dots & \dots & \dots & \dots \\ p_{n1}(k) & p_{n2}(k) & \dots & p_{nn}(k) \end{pmatrix}.$$

Для неоднородной цепи Маркова вектор распределения вероятностей состояний определяется как

$$(p_1(k), p_2(k), \dots, p_n(k)) = (p_1(k-1), p_2(k-1), \dots, p_n(k-1))\mathcal{P}^{(k)} \quad (12.2.3)$$

Так же для неоднородной цепи имеет место следующая формула:

$$(p_1(k), p_2(k), \dots, p_n(k)) = (p_1(0), p_2(0), \dots, p_n(0))\mathcal{P}^{(1)}\mathcal{P}^{(2)} \dots \mathcal{P}^{(k)} \quad (12.2.4)$$

Вероятности состояний $p_i(k)$ $i = 1, 2, \dots$ неоднородной цепи Маркова на каждом шаге вычисляются либо по рекуррентной формуле (12.2.3), либо по формуле (12.2.4), где $(p_1(0), p_2(0), \dots, p_n(0))$ — вектор начального распределения вероятностей состояний системы.

12.3. Примеры цепей Маркова

Большое число физических, биологических и экономических явлений описываются цепями Маркова.

♦ **Пример 1. Пространственно-однородные цепи Маркова**

Пусть дискретная случайная величина ξ принимает неотрицательные целочисленные значения, причем $P\{\xi = i\} = p_i$, и $p_i \geq 0$ и $\sum_{i=1}^{\infty} p_i = 1$. Пусть $\xi_1, \xi_2, \dots, \xi_n, \dots$ — результаты независимых наблюдений случайной величины ξ .

а) Определим процесс $\{X_n\}$, $n = 0, 1, 2, \dots$, положив

$X_n = \xi_n$, где $X_0 = \xi_0$ задано. Матрица переходных вероятностей этого процесса имеет вид

$$\mathcal{P} = \begin{pmatrix} p_0 & p_1 & \dots & p_n \\ p_0 & p_1 & \dots & p_n \\ \dots & \dots & \dots & \dots \\ p_0 & p_1 & \dots & p_n \end{pmatrix}.$$

У \mathcal{P} все строки одинаковы, следовательно, случайная величина X_{n+1} не зависит от случайной величины X_n .

б) Определим процесс $\{X_n\}$, $n = 0, 1, 2, \dots$, положив $X_n = S_n$, где $S_n = \xi_1 + \xi_2 + \dots + \xi_n$. Считаем, что $S_0 = 0$. Найдем матрицу переходных вероятностей.

$$P\{X_{n+1} = j / X_n = i\} = P\{\xi_1 + \dots + \xi_{n+1} = j / \xi_1 + \dots + \xi_n\} = P\{\xi_{n+1}\} =$$

$$= \begin{cases} p_{j-i}, & j \geq i, \\ 0, & j < i \end{cases}.$$

$$\mathcal{P} = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \dots \\ \dots & p_0 & p_1 & p_2 & \dots \\ \dots & \dots & p_0 & p_1 & \dots \\ \dots & \dots & \dots & p_1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Процесс $\{X_n\}$ образует ЦМ.

◇ **Пример 2. Одномерные случайные блуждания**

Пусть η_1, η_2, \dots — независимые одинаково распределенные случайные величины. $P\{\eta_k = 1\} = p$, $P\{\eta_k = -1\} = q$, а последовательность $\{\xi_t\}$ строится по правилу

$$\xi_t = \max\{\xi_{t-1} + \eta_t, 0\}, \quad t = 0, 1, 2, \dots$$

и называется случайным блужданием на множестве неотрицательных целых чисел с отражающим экраном в 0.

$$P\{\xi_{t+1} = i + 1 / \xi_t = i\} = p,$$

$$P\{\xi_{t+1} = i - 1 / \xi_t = i\} = q = 1 - p, \quad i \geq 1,$$

$$P\{\xi_{t+1} = 0 / \xi_t = 0\} = q.$$

$$\mathcal{P} = \begin{pmatrix} q & p & 0 & 0 & \dots \\ q & 0 & p & 0 & \dots \\ 0 & q & 0 & p & \dots \\ 0 & 0 & q & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

◇ **Пример 3. Модель Эренфестов для процесса диффузии**

Эта модель представляет собой цепь с $n + 1$ состояниями и возможными переходами только в состояния, соседние справа и слева. Тогда переходные вероятности определяются равенствами:

$$p_{k,k+1} = 1 - \frac{k}{n}, \quad p_{k,k-1} = \frac{k}{n}.$$

Цепь имеет две физические интерпретации. Подробно рассмотрим первую модель, названную по имени Пауля и Татьяны Эренфестов. Ученые физики описали мысленный эксперимент, при котором перемещали n частиц по двум сосудам. В каждый момент времени $t = 0, 1, \dots$ случайно, равновероятно и независимо от предыстории выбирается одна из n частиц и перемещается в другой сосуд. Пусть ξ_t — число частиц в первом сосуде в момент

t . Состоянием процесса считается число частиц $(0, 1, 2, \dots, n)$ в первом сосуде, и переходные вероятности имеют вид:

$$P\{\xi_{t+1} = k - 1 / \xi_t = k\} = \frac{k}{n}, \quad P\{\xi_{t+1} = k + 1 / \xi_t = k\} = 1 - \frac{k}{n}.$$

Тогда

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ \frac{1}{n} & 0 & 1 - \frac{1}{n} & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{n} & 0 & 1 - \frac{2}{n} & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 - \frac{2}{n} & 0 & \frac{2}{n} & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 - \frac{1}{n} & 0 & \frac{1}{n} \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Следовательно, эксперимент Эренфестов описывается цепью Маркова. Со второй интерпретацией, представляющей случайное блуждание, в которой вероятность перемещения изменяется при изменении положения (диффузия при наличии центральной силы), читатель подробно может познакомиться в книге [18].

♦ Пример 4. Ветвящиеся процессы⁷

Предположим, что организм в конце своего времени жизни производит случайное число ξ потомков согласно распределению вероятностей

$$P\{\xi = k\} = p_k, \quad k = 0, 1, 2, \dots,$$

⁷В 1873–1874 г. Гальтон и де Кандоль указали на ряд примеров исчезновения фамилий английских лордов. Ими возможно впервые был поставлен вопрос: какова вероятность исчезновения популяции с течением времени? Модель такой ситуации была предложена и частично решена священником Ватсоном в 1874 г. Полный анализ модели, которая сегодня называется ветвящимся процессом, был закончен Стефансом в 1930 г. Сам же термин "ветвящиеся процессы" был введен А.Н. Колмогоровым. Бурное развитие теории ветвящихся процессов началось с известных работ А.Н. Колмогорова — "Ветвящиеся случайные процессы", "Вычисление финальных вероятностей", выполненных совместно с его учениками Н.А. Дмитриевым и Б.А. Севастьяновым.

где $p_k \geq 0$, $\sum_{k=0}^{\infty} p_k = 1$. В свою очередь потомки независимо друг от друга в конце своего времени жизни производят потомство, согласно того же распределения. Процесс $\{X_n\}$ — марковская цепь, где X_n — численность популяции в n -ом поколении. Действительно, если задано значение, например, $X_0 = 1$, то матрица переходных вероятностей определяется

$$p_{ij} = P\{X_{n+1} = j / X_n = i\} = P\{\xi_1 + \xi_2 + \dots + \xi_i = j\},$$

где $\xi_1 + \xi_2 + \dots + \xi_i$ — общее число потомков.

Подобные процессы служат не только для математического описания роста популяции, но и других явлений, связанных с ветвлением, например, распространение инфекции, химические или ядерные реакции и др..

12.4. Классификация состояний цепей Маркова

Будем рассматривать только однородные цепи Маркова. Пусть E — множество состояний цепи Маркова, $p_{ij} = P\{X_n = j / X_0 = i\}$ — вероятность перехода за n шагов из состояний i в состояние j ; $f_{ii}^{(n)} = P\{X_n = i, X_{n-1} \neq i, \dots, X_1 \neq i / X_0 = i\}$ — вероятность первого возвращения за n шагов в состояние i .

Определение 12.4.1. Состояние $i \in E$ называется *несущественным*, если найдется $j \in E$, такое что $p_{ij}^{(m)} > 0$ для некоторого $m \geq 1$, но $p_{ji}^{(n)} = 0$ для всех $n \geq 1$. В противном случае состояние i называется *существенным*.

♦ **Пример 1.** В цепи Маркова с матрицей переходных вероятностей

$$\mathcal{P} = \begin{pmatrix} \alpha & \beta & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\alpha + \beta + \gamma = 1, \alpha, \beta, \gamma > 0$$

состояние e_1 несущественно: $p_{11}^{(n)} = \alpha^n \rightarrow 0$;

$$p_{12}^{(n)} = (1 - p_{11}^{(n)}) \frac{\beta}{\beta + \gamma}; \quad p_{13}^{(n)} = (1 - p_{11}^{(n)}) \frac{\gamma}{\beta + \gamma}.$$

Определение 12.4.2. Состояние i называется поглощающим, если $p_{ii} = 1$.

В примере 1 состояния e_2 и e_3 - поглощающие.

Определение 12.4.3. Состояния $i, j \in E$ называются сообщающимися, если найдутся такие $m, n \geq 1$, что $p_{ij}^{(m)} > 0$ и $p_{ji}^{(n)} > 0$.

Обозначение $i \leftrightarrow j$.

Свойства:

- 1) рефлексивность: $i \leftrightarrow i$;
- 2) симметричность: если $i \leftrightarrow j$, то $j \leftrightarrow i$;
- 3) транзитивность: если $i \leftrightarrow j$ и $j \leftrightarrow k$, то $i \leftrightarrow k$.

Из свойств 1–3 следует, что все множество состояний E можно разбить на классы эквивалентности. Состояния объединяются в один класс, если они сообщаются друг с другом.

Определение 12.4.4. Цепь Маркова, состоящая из одного класса сообщающихся состояний, называется неразложимой. Если она содержит более одного класса сообщающихся состояний, то она разложима.

♦ **Пример 2.** Пусть цепь Маркова имеет следующую матрицу переходных вероятностей:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \vdots & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & 0 & 1 & 0 \\ 0 & 0 & \vdots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \vdots & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}.$$

Состояния цепи Маркова распадаются на 2 класса сообщающихся состояний: $\{e_1, e_2\}$ и $\{e_3, e_4, e_5\}$. В зависимости от начального состояния процесс разворачивается либо только в первом классе состояний и его переходы описываются подматрицей P_1 , либо только во втором классе и его переходы описываются подматрицей P_2 .

Определение 12.4.5. Пусть d_i — НОД чисел

$$\{n \geq 1 : p_{ii}^n > 0\}.$$

Состояние i называется периодическим с периодом d_i , если

$$d_i > 1.$$

В противном случае состояние — аperiodическое.

◇ **Пример 3.** Пусть

$$\mathcal{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

$$\text{Тогда } \mathcal{P}^{(2)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{P}^{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{P}^4 = \mathcal{P}$$

и т.д.

Следовательно,

$$p_{i,j}^{(n)} = \begin{cases} 1, & \text{если } n \equiv j - i \pmod{3}; \\ 0, & \text{в остальных случаях} \end{cases}.$$

Примем без доказательства следующие утверждения:

1. Если $i \longleftrightarrow j$, то $d(i) = d(j)$.

Это утверждение определяет период как характеристику класса сообщающихся состояний.

2. Если состояние i имеет период $d(i)$, то $\exists N < \infty$ и $p_{ij}^{(nd(i))} > 0$ для всех $n > N$.

Этим утверждается, что возвращение в состояние i может происходить во все достаточно далёкие моменты времени, кратные периоду $d(i)$.

Определение 12.4.6. Существенное состояние $i \in E$ называется возвратным, если $\exists n < \infty$ такое что $P\{X_n = i / X_0 = i\} = 1$, и невозвратным, если $P\{X_n = i / X_0 = i\} < 1$.

Определение 12.4.7. Существенное состояние $i \in E$ называется возвратным нулевым, если $p_{ii}^{(n)} \rightarrow 0$ при $n \rightarrow \infty$.

Определение 12.4.8. *Неразложимая цепь Маркова называется аperiодической, если все её состояния аperiодические.*

Для неразложимой цепи Маркова справедливы свойства:

а) если хотя бы одно состояние возвратно, то и все другие возвратны;

б) если хотя бы одно состояние нулевое, то все другие состояния нулевые;

в) если хотя бы одно состояние имеет период $d > 1$, то и все остальные периодичны с периодом d .

Таким образом, неразложимая цепь Маркова будет аperiодической, если хотя бы одно из её состояний — аperiодическое.

Теорема 12.4.1 (критерий возвратности). *Состояние i является возвратным тогда и только тогда, когда*

$$p_i = \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty.$$

Доказательство. Пусть $\nu_n = f_{ii}^{(n)}$ — вероятность впервые вернуться из состояния i в состояние i через n шагов, тогда $\omega_n = \sum_{m=0}^n f_{ii}^{(m)} f_{ii}^{(n-m)}$ — вероятность любым способом вернуться в состояние i через n шагов (вначале через m шагов, а затем через оставшиеся $(n-m)$ шагов). С учётом введенных обозначений по формуле полной вероятности можно записать:

$$\omega_n = \nu_0 \omega_n + \nu_1 \omega_{n-1} + \dots + \nu_n \omega_0,$$

и пусть $\nu_0 = 0$, $\omega_0 = 1$. Обратимся к производящим функциям:

$$V(z) = \sum_{m=0}^{\infty} \nu_m z^m, \quad W(z) = \sum_{n=0}^{\infty} \omega_n z^n, \quad 0 \leq z \leq 1.$$

Полученные соотношения между коэффициентами выражают равенство

$$W(z) - \omega_0 = W(z)V(z), \quad \omega_0 = 1.$$

Отсюда $W(z) = 1/(1 - V(z))$. По определению возвратности состояния i следует, что

$$\nu = \sum_{m=0}^{\infty} \nu_m = \lim_{z \rightarrow 1} V(z) = 1.$$

Тогда

$$\lim_{z \rightarrow 1} W(z) = \lim_{z \rightarrow 1} \frac{1}{1 - V(z)} = \infty.$$

Но

$$\lim_{z \rightarrow 1} W(z) = \lim_{z \rightarrow 1} \sum_{n=0}^{\infty} \omega_n z^n = \sum_{n=0}^{\infty} \omega_n$$

и, таким образом, возвратность состояния i равносильна тому, что ряд $\sum_{n=0}^{\infty} \omega_n$ расходится, где $\omega_n = \sum_{i=1}^{\infty} p_{ii}^{(n)}$.

Доказательство (12.4.1) закончено.

◇ **Пример 4.** Рассмотрим одномерное случайное блуждание по целочисленной решётке $i = 0, \pm 1, \pm 2, \dots$. За каждый период частица с вероятностью p перемещается на единицу вправо и с вероятностью q — на единицу влево. Очевидно, что $p_{ii}^{(2n+1)} = 0$, а

$$p_{ii}^{(2n)} = C_{2n}^n p^n q^n = \frac{(2n)!}{n!n!} p^n q^n.$$

Используя формулу Стирлинга $n! \sim \sqrt{2\pi n} n^n e^{-n}$, получаем:

$$p_{ii}^{(2n)} = \frac{(pq)^n 2^{2n}}{\sqrt{\pi n}} \rightarrow 0 \text{ при } n \rightarrow \infty,$$

то есть все состояния нулевые. Применяя критерии возвратности, заключаем, что при симметричном случайном блуждании, т.е. когда $p = q = \frac{1}{2}$, то $p_{ii}^{(2n)} \sim \frac{1}{\sqrt{\pi n}}$ и ряд $\sum_{n=1}^{\infty} p_{ii}^{(2n)} = \infty$, все состояния возвратны. Таким образом, одномерное случайное блуждание возвратно тогда и только тогда, когда $p = q = \frac{1}{2}$. Если $p \neq q$, то $\sum_{n=1}^{\infty} p_{ii}^{(2n)} < \infty$, то каждое состояние i является невозвратным.

Теорема 12.4.2. *Если исходное состояние i является возвратным, то система с вероятностью 1 за бесконечно много*

число шагов бесконечно много раз возвращается в i . Если это состояние является невозвратным, то за бесконечное число шагов система с вероятностью 1 лишь конечное число раз бывает в состоянии i .

12.5. Предельная теорема для конечных цепей Маркова

Существенной характеристикой случайных процессов является их поведение при $n \rightarrow \infty$. Рассмотрим однородную цепь Маркова.

Теорема 12.5.1 (Теорема Маркова)⁸. Если при некотором $n < \infty$ все элементы матрицы положительны, то существуют такие постоянные числа p_j , $j \in \{1, 2, \dots, n\}$, что независимо от индексов i имеют место равенства $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = p_j$, где $\sum_{j=1}^n p_j = 1$, $p_j > 0$.

Физический смысл теоремы следующий: вероятность системы находится в состоянии j , практически не зависит от того, в каком состоянии она находилась в далёком прошлом.

Определение 12.5.1. Распределение вероятностей p_j^* , $j = 0, \pm 1, \dots$, $\sum_j p_j^* = 1$ называют стационарным для однородной цепи Маркова X_n , $n \geq 0$, если при заданном начальном распределении в последующие моменты времени распределение вероятностей $p_j(n) = P\{X_n = j\}$ остаётся неизменным при всех $n \geq 0$: $p_j(n) = p_j^*$ или в общем виде $p_j^* = p_j^* P$.

Определение 12.5.2. Цепь Маркова, удовлетворяющая условиям теоремы Маркова, называется эргодической⁹, а распределение вероятностей $p^* = (p_0, p_1, \dots)$ — стационарным распределением цепи Маркова.

Примем без доказательства следующую теорему.

⁸Следует отметить, что с доказательством теоремы читатель может ознакомиться в книге [4].

⁹Синонимом термина "эргодическое состояние" является "возвратное, ненулевое, апериодическое состояние". Иногда эргодическим называют любое возвратное состояние (см. [18].)

Теорема 12.5.2. *Для того чтобы конечная цепь Маркова была эргодической, необходимо и достаточно, чтобы она была неразложимой и апериодичной.*

§ 13. Дискретный случайный марковский процесс

В 1931 г. была опубликована статья А.Н. Колмогорова "Об аналитических методах в теории вероятностей", где были заложены основы теории марковских процессов. Получены уравнения (прямые и обратные), которые управляют вероятностями перехода. Общая теория марковских процессов была создана в 30–40 гг. работами А.Н. Колмогорова, В. Феллера, В. Деблина, П. Леви и др.

13.1. Вероятностные характеристики марковских процессов

Пусть $\{X(t)\}$, $t \geq 0$ — случайный процесс, принимающий значения при каждом t из множества E . Далее без ограничения общности будем считать, что $i_k = k$, $k = 0, 1, \dots$

Определение 13.1.1. *Случайная функция $\{X(t)\}$, $t \geq 0$, принимающая значения из множества E , называется марковским процессом с непрерывным временем и дискретным множеством значений, если для любых элементов $0 \leq t_1 \leq t_2 \leq \dots \leq t_{n-1} \leq s \leq t$ и значений $i_1, i_2, \dots, i_{n-1}, i, j \in E$ выполнено*

$$\begin{aligned} P\{X(t) = j / X(s) = i, X(t_{n-1}) = i_{n-1}, \dots, X(t_1) = i_1\} = \\ = P\{X(t) = j / X(s) = i.\} \end{aligned}$$

Определение 13.1.2. *Вероятностью перехода марковского процесса $\{X(t)\}$ называется функция*

$$p_{i,j}(s, t) = P\{X(t) = j / X(s) = i\}, \quad \text{где } i, j \in E, 0 \leq s \leq t.$$

Замечание. Из определения следуют свойства вероятностей перехода:

1. $\sum_{j \in E} p_{ij}(s, t) = 1$ для всех $i \in E$.
2. $p_{ii}(t, t) = 1$ для всех $i \in E$.
3. $p_{ij}(t, t) = 0$ для всех $i \neq j$.

Возможные значения процесса $\{X(t)\}$, то есть элементы множества E , называют также состояниями процесса $\{X(t)\}$. Поэтому $p_{ij}(s, t)$ — вероятность перехода из состояния i в момент s в состояние j в момент t , $i, j \in E$.

Определение 13.1.3. Вероятностью i -того состояния в момент времени $t \geq 0$ называется величина

$$p_i(t) = P\{X(t) = i\}, \quad \text{где } i \in E.$$

Очевидно, что

$$p_i(t) \geq 0, \quad \sum_{i \in E} p_i(t) = 1.$$

Между переходными вероятностями и вероятностями состояний имеется связь.

Теорема 13.1.1. Пусть $0 \leq s \leq u \leq t$, тогда

$$p_{ij}(s, t) = \sum_k p_{ik}(s, u) p_{kj}(u, t),$$

$$p_i(t) = \sum_k p_k(s) p_{ki}(s, t).$$

Доказательство. Доказывается по формуле полной вероятности.

Замечание. Равенства, сформулированные в теореме (13.1.1), называются уравнениями Колмогорова-Чепмена.

Определение 13.1.4. Марковский процесс $\{X(t)\}$ называется однородным, если $p_{ij}(s, s+t) = p_{ij}(0, t)$, для всех $i, j \in E$, $s, t \geq 0$.

Предположим, что при каждом $s \geq 0$ переходная плотность дифференцируема по t при любых $i, j \in E$.

Определение 13.1.5. *Интенсивностью (плотностью вероятности) перехода $\lambda_{ij}(t) \geq 0$ из состояния i в состояние j в момент $t \geq 0$ называется величина*

$$\lambda_{ij}(t) = \lim_{h \rightarrow 0} \frac{p_{ij}(t, t+h)}{h}, \quad i \neq j.$$

◇ **Пример 1.** Показать, что интенсивности переходов однородного процесса не зависят от времени.

Решение. В силу однородности $p_{ij}(t, t+h) = p_{ij}(0, h)$, поэтому

$$\lambda_{ij}(t) = \lim_{h \rightarrow 0} \frac{p_{ij}(0, h)}{h} = \lambda_{ij}(0) = \text{const.}$$

Далее мы будем рассматривать только однородные процессы, вследствие этого

$$\lambda_{ij}(t) = \lambda_{ij}, \quad p_{ij}(s, s+t) = p_{ij}(t).$$

13.2. Система дифференциальных уравнений Колмогорова

Наибольший практический интерес при исследовании процессов описанного типа представляет вычисление вероятностей его состояния в любой момент $t \geq 0$, если задано начальное распределение вероятностей состояния $p(0)$.

Теорема 13.2.1. *Пусть $\{X(t)\}$ марковский процесс, причем для всех пар состояний i и j определены плотности вероятностей λ_{ij} и λ_{ji} . Тогда вероятности состояний системы $p_i(t)$ удовлетворяют системе уравнений Колмогорова:*

$$\begin{cases} \frac{\partial p_i}{\partial t} = -p_i(t) \sum_{k \neq i}^n \lambda_{ik} + \sum_{k \neq i}^n p_k(t) \lambda_{ki} \\ \sum_{i=1}^n p_i(t) = 1. \end{cases}$$

Доказательство.

Используем равенство Колмогорова-Чепмена, а именно

$$p_i(t) = \sum_k p_k(s) p_{ki}(s, t) \quad (13.2.1)$$

и используем определение интенсивности переходов, тогда

$$p_{ki}(h) = \lambda_{ki}h + o(h), \quad k \neq i.$$

Учитывая, что $\sum_k p_{ik}(h) = 1$, получим

$$p_{ii}(h) = 1 - \sum_{\substack{k=1 \\ k \neq i}}^n p_{ik}(h) = 1 - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik}h + o(h).$$

Подставляя последние два условия в равенство (13.2.1), имеем:

$$p_i(t+h) = p_i(t) \left(1 - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik}h\right) + \sum_{\substack{k=1 \\ k \neq i}}^n p_k(t) \lambda_{ki}h + o(h)$$

или, что то же самое,

$$\frac{p_i(t+h) - p_i(t)}{h} = - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} p_i(t) + \sum_{\substack{k=1 \\ k \neq i}}^n p_k(t) \lambda_{ki} + \frac{o(h)}{h}.$$

Для завершения доказательства достаточно перейти к пределу при $h \rightarrow 0$

$$\frac{\partial p_i}{\partial t} = -p_i(t) \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} + \sum_{\substack{k=1 \\ k \neq i}}^n p_k(t) \lambda_{ki}, \quad i \in E,$$

или

$$\begin{cases} \frac{\partial p_i}{\partial t} = -p_i(t) \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} + \sum_{\substack{k=1 \\ k \neq i}}^n p_k(t) \lambda_{ki} \\ \sum_{i=1}^n p_i(t) = 1. \end{cases} \quad (13.2.2)$$

Доказательство (13.2.1) закончено.

Полученные уравнения (13.2.2) называются дифференциальными уравнениями Колмогорова.

Замечание. Система (13.2.2) не является линейно независимой, т.е. она является избыточной и имеет единственное решение при каждом $p(0) = \{p_0(0), p_1(0), \dots\}$.

Для практического составления уравнений Колмогорова удобно пользоваться графическим представлением процесса в виде стохастического графа.

Правило.

Вершинами графа являются состояния процесса. Стрелками указываются возможные переходы, а рядом с каждой стрелкой записывается соответствующая интенсивность перехода. В левой части каждого уравнения стоит дробь

$\frac{\partial p_k}{\partial t}$, в правой части столько слагаемых, сколько стрелок связано с состоянием i_k . Слагаемые берутся со знаком "плюс", если стрелка направлена к состоянию, и со знаком "минус", если стрелка выходит из состояния i_k .

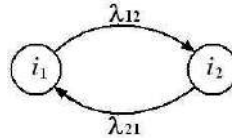


Рис. 21

Каждое слагаемое равно произведению интенсивности перехода, отмеченной на стрелке, на вероятность того состояния, из которого стрелка выходит (рис. 21).

$$\begin{cases} \frac{\partial p_1}{\partial t} = -\lambda_{12}p_1 + \lambda_{21}p_2 \\ \frac{\partial p_2}{\partial t} = -\lambda_{21}p_2 + \lambda_{12}p_1 \\ p_1 + p_2 = 1. \end{cases}$$

♦ **Пример 1.** Двухпроцессорная вычислительная система предназначена для обработки простейшего потока задач, поступающих с интенсивностью λ . Производительность процессоров одинакова. Задача случайным образом принимается на обслуживание процессором. Если оба процессора заняты, пользователь получает отказ. Постройте граф функционирования данной системы и запишите систему дифференциальных уравнений Колмогорова по размеченному графу, если интенсивность решения задач первым процессором равна μ_1 , вторым — μ_2 .

Решение. Рассмотрим возможные состояния системы, которые определяются состояниями процессоров: i_0 — оба процессора простаивают; i_1 — первый процессор занят решением задач, второй простаивает; i_2 — второй процессор занят, первый простаивает; i_3 — оба процессора заняты решением задач.

Граф функционирования системы имеет вид (рис. 22):

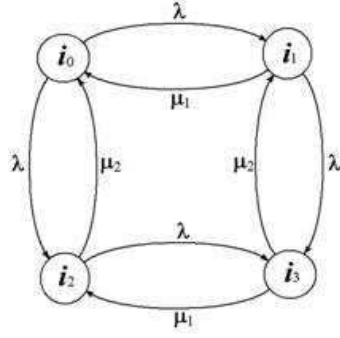


Рис. 22

По графу запишем систему линейных дифференциальных уравнений А.Н. Колмогорова:

$$\begin{cases} \frac{\partial p_0}{\partial t} = -2\lambda p_0 + \mu_2 p_2 + \mu_1 p_1 \\ \frac{\partial p_1}{\partial t} = \mu_2 p_3 + \lambda p_0 - \lambda p_1 - \mu_1 p_1 \\ \frac{\partial p_2}{\partial t} = \lambda p_1 + \lambda p_2 - \mu_2 p_3 - \mu_1 p_3 \\ \frac{\partial p_3}{\partial t} = \mu_1 p_3 + \lambda p_0 - \mu_2 p_2 - \lambda p_2 \\ \sum_{i=0}^3 p_i = 1, \end{cases}$$

или

$$\begin{cases} \frac{\partial p_0}{\partial t} = -2\lambda p_0 + \mu_1 p_1 + \mu_2 p_2 \\ \frac{\partial p_1}{\partial t} = \lambda p_0 - p_1(\lambda + \mu_1) + \mu_2 p_3 \\ \frac{\partial p_2}{\partial t} = \lambda p_1 + \lambda p_2 - p_3(\mu_2 + \mu_1) \\ \frac{\partial p_3}{\partial t} = \lambda p_0 - p_2(\mu_2 + \lambda) + \mu_1 p_3 \\ \sum_{i=0}^3 p_i = 1. \end{cases}$$

13.3. Эргодические свойства однородных марковских случайных процессов

Определение 13.3.1. Марковский процесс $\{X(t)\}$, $t \geq 0$ называется эргодическим, если для любого начального распределения вероятностей состояний $p_k(0)$, $k \in E$ предельные вероятности $\lim_{t \rightarrow \infty} p_k(t) = \tilde{p}_k$, $k \in E$, причем $\{\tilde{p}_k, k \in E\}$ не зависят от $\{p_k(0), k \in E\}$ и удовлетворяют условию $\tilde{p}_k \geq 0$, $\sum_k \tilde{p}_k = 1$.

При этом распределение вероятностей $\{\tilde{p}_k, k \in E\}$ называется стационарным распределением процесса $X(t)$.

Если выполнены условия, указанные в определении (13.3.1), то распределение $\{\tilde{p}_k, k \in E\}$ существует и удовлетворяет алгебраическим уравнениям Колмогорова:

$$\begin{cases} -\tilde{p}_i(t) \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} + \sum_{\substack{k=1 \\ k \neq i}}^n \tilde{p}_k(t) \lambda_{ki} = 0 \\ \sum_{i=1}^n \tilde{p}_i(t) = 1, \end{cases} \quad (13.3.1)$$

причем решение системы единственно.

Замечание 13.3.1. Система уравнений (13.3.1) получается из системы дифференциальных уравнений (13.2.2), если положить $\frac{\partial p_i}{\partial t} = 0$, $p_i(t) = \tilde{p}_i$, $i \in E$.

Замечание 13.3.2. Система (13.3.1) справедлива, если $\{X(t)\}$ имеет конечное число сообщающихся состояний.

♦ **Пример 1.** Рассмотрим простейшую задачу теории массового обслуживания — задачу о функционировании одноканальной системы обслуживания с отказами, на вход которой поступает простейший поток заявок с интенсивностью λ (заявка, заставшая канал занятым, покидает систему), а интенсивность обслуживания заявки равна величине μ . Определим предельные вероятности состояний рассматриваемого процесса.

Решение. В данном случае система имеет два возможных состояния: i_1 — канал свободен; i_2 — канал занят (рис. 23). Математическая модель данного процесса массового обслуживания имеет вид:

$$\begin{cases} \frac{\partial p_1}{\partial t} = -\lambda p_1 + \mu p_2 \\ \frac{\partial p_2}{\partial t} = -\mu p_2 + \lambda p_1 \\ p_1 + p_2 = 1, \end{cases}$$

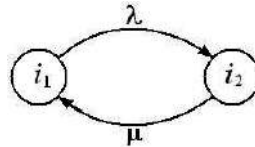


Рис. 23

Зададим начальное распределение $p(0) = \{\alpha, 1 - \alpha\}$. Подставляя $p_1 = 1 - p_2$ в первое уравнение системы, получаем:

$$\begin{cases} \frac{\partial p_1}{\partial t} = -\lambda p_1 + \mu(1 - p_1) \\ p_1(0) = \alpha. \end{cases}$$

Решим задачу Коши для линейного неоднородного уравнения

$$p_1(t) = \left(\alpha - \frac{\mu}{\mu + \lambda} \right) e^{-(\lambda + \mu)t} + \frac{\mu}{\mu + \lambda}.$$

В силу того, что $\mu + \lambda > 0$, $e^{-(\lambda + \mu)t} \rightarrow 0$ при $t \rightarrow \infty$, получаем по определению

$$\tilde{p}_1 = \lim_{t \rightarrow \infty} p_1(t) = \frac{\mu}{\mu + \lambda}.$$

Аналогично получаем:

$$\tilde{p}_2 = \lim_{t \rightarrow \infty} p_2(t) = \frac{\lambda}{\mu + \lambda}.$$

Итак,

$$\tilde{p} = \left(\frac{\mu}{\mu + \lambda}; \frac{\lambda}{\mu + \lambda} \right).$$

Замечание 13.3.3. Полученные результаты можно вывести, если воспользоваться не определением, а системой дифференциальных уравнений (13.2.2):

$$\begin{cases} -\lambda \tilde{p}_1 + \mu \tilde{p}_2 = 0 \\ -\mu \tilde{p}_2 + \lambda \tilde{p}_1 = 0 \\ \tilde{p}_1 + \tilde{p}_2 = 1 \end{cases}$$

Решив систему, получим те же результаты (рис. 24).

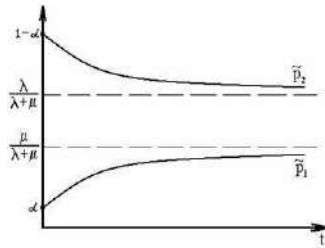


Рис. 24

13.4. Пуассоновский процесс

Пусть $\{X(t)\}$, $0 \leq t < \infty$ — марковский процесс, принимающий неотрицательные целочисленные значения. Обозначим $X(t)$ — число появления события за время t , $P(t; t+h)$ — вероятность появления хотя бы одного события на промежутке времени $[t; t+h]$. Предположим, что при каждом t существует постоянная $\lambda(t) = \lim_{h \rightarrow 0} \frac{P(t, t+h)}{h}$, которая называется интенсивностью процесса.

Определение 13.4.1. *Простейшим потоком событий называется целочисленный марковский процесс, удовлетворяющий следующим условиям:*

- 1) $\{X(t)\}$ — однородный или стационарный процесс, то есть его приращение зависит от величины приращения аргумента $P\{X(t+h) - X(t) = k\} = P\{X(h) = k\}$, $\forall t, h \geq 0$.
- 2) $\{X(t)\}$ — ординарный процесс, то есть события в потоке следуют строго одно за другим и не происходят вместе

$$P\{X(t+h) - X(t) = 1\} = \lambda h + o(h),$$

$$P\{X(t+h) - X(t) > 1\} = o(h).$$

- 3) $\{X(t)\}$ — процесс без последствий, то есть случайные величины $\{X(t_1), X(t_2) - X(t_1), \dots\}$ независимы в совокупности.

Замечание 13.4.1. Ординарный процесс без последствий называется пуассоновским процессом.

Из свойства 2 получим два важных следствия:

$$\text{а) } P\{X(t+h) - X(t) = 0\} = 1 - P\{X(t+h) - X(t) \geq 1\} = 1 - \lambda h + o(h).$$

$$\text{б) } P\{X(0) = 0\} = 1.$$

Действительно, $P\{X(t+h) - X(t) = 0\} = P\{X(h) = 0\} = 1 - \lambda h + o(h)$, если $h = 0$, то $P\{X(0) = 0\} = 1$.

Условие 2, 2а, 2б с учетом условия 3 в обычных обозначениях можно записать в виде:

$$p_{ik}(h) = \begin{cases} 0, & i > k \\ \lambda h + 0(h), & k - i = 1 \\ 1 - \lambda h + 0(h), & k - i = 0 \\ 0(h), & k - i > 1. \end{cases}$$

В соответствии с условием граф состояний имеет вид (рис. 25):

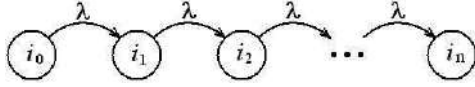


Рис. 25

Составим по графу уравнения Колмогорова

$$\begin{cases} \frac{\partial p_0}{\partial t} = -\lambda p_0 \\ \frac{\partial p_1}{\partial t} = -\lambda p_1 + \lambda p_0 \\ \dots\dots\dots \\ \frac{\partial p_k}{\partial t} = -\lambda p_k + \lambda p_{k-1} \\ \sum_k p_k(t) = 1. \end{cases}$$

Начальные условия $p(0) = \{1, 0, 0, \dots, 0\}$ имеют вид, обусловленный следствием (б).

Для решения системы уравнений Колмогорова рассмотрим производящую функцию $\varphi(z, t)$, $|z| \leq 1$.

$$\varphi(z, t) = Ez^{W(t)} = \sum_{k=0}^{\infty} z^k p_k(t), \quad |z| \leq 1.$$

Очевидно, что $\varphi(z, t)$ удовлетворяет условию уравнения

$$\begin{cases} \frac{\partial \varphi}{\partial t} = -\lambda \varphi + z \lambda \varphi \\ \varphi(z, 0) = z^0 p_0(0) + z p_1(0) + \dots = 1. \end{cases}$$

Получим:

$$\varphi(z, t) = ce^{-(\lambda - \lambda z)t},$$

$$\varphi(z, 0) = c = 1,$$

$$\varphi(z, t) = e^{-(\lambda - \lambda z)t} = e^{-\lambda t} \left(1 + \lambda z t + z^k \frac{(\lambda t)^k}{k!} + \dots \right),$$

следовательно, $p_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$.

В полученном выражении $X(t) = \lambda t$ — среднее число появления события за время $(0; t)$.

Таким образом, при каждом t случайная величина $X(t)$ распределена по закону Пуассона с параметром λt .

Замечание 13.4.2. Процесс Пуассона не имеет ни стационарного, ни предельного распределения, так как

$$\lim_{t \rightarrow \infty} p_k(t) = \tilde{p}_k = 0, \forall k, \quad \text{и следовательно} \quad \sum_{k=0}^{\infty} \tilde{p}_k \neq 1.$$

Один и тот же математический объект можно интерпретировать по-разному. Пуассоновский процесс мы определили как ординарный процесс с независимыми приращениями. Если же элементы скачков процесса (пуассоновский процесс принимает только целые значения и его траектории — монотонно неубывающие ступенчатые функции) отождествить с точками на действительной оси, то можно будет считать, что пуассоновский процесс — это совокупность случайных точек на прямой или поток моментов наступления событий на оси времени. Поясним сказанное. Пусть система подвержена мгновенным изменениям, которые могут быть обусловлены такими событиями, как распад физической частицы, телефонный вызов и др.. Все изменения подобны друг другу, и мы интересуемся только их общим числом. Каждое такое изменение и фиксируется на временной оси, нас же интересует случайное распределение точек на действительной оси.

Рассмотрим распределение промежутков между моментами событий в пуассоновском потоке с интенсивностью $\lambda > 0$. Пусть τ_k — случайная величина, характеризующая длину интервала времени, в течение которого процесс изменяет свое состояние с " $k-1$ " на " k "-ое.

Теорема 13.4.1 . Если $\{X(t)\}$ — однородный пуассоновский процесс с интенсивностью λ , то распределение промежутков времени τ_k между моментами его скачков независимо и имеет

показательное распределение с функцией распределения

$$F_{\tau_k}(t) = \begin{cases} 1 - e^{-\lambda t}, & t > 0 \\ 0, & t \leq 0. \end{cases}$$

Доказательство.

$$F_{\tau_k}(t) = P\{\tau_k < t\} = 1 - P\{\tau_k \geq t\},$$

так как $\tau_k = t_k - t_{k-1}$, то при $\tau_k \geq t$ получим:

$$P\{\tau_k \geq t\} = P\{X(t_{k-1} + t) - X(t_{k-1}) = 0\} = P\{X(t) = 0\} = e^{-\lambda t}.$$

Следовательно,

$$F_{\tau_k}(t) = 1 - P\{\tau_k \geq t\} = \begin{cases} 1 - e^{-\lambda t}, & t > 0 \\ 0, & t \leq 0. \end{cases}$$

Доказательство (13.4.1) закончено.

Независимость случайной величины τ_k от поведения пуассоновского процесса $\{X(t)\}$ на полуоси $(0; \infty)$ следует из независимости приращений $X(t)$ на непересекающихся интервалах. Дифференцируя по t , легко находим плотность вероятности

$$p_{\tau_k}(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0 \\ 0, & t \leq 0. \end{cases}$$

Пусть случайная величина t_k — момент времени, в который процесс из " $k-1$ " состояния переходит в " k "-ое состояние.

Теорема 13.4.2 Если $\{X(t)\}$ — однородный пуассоновский процесс с интенсивностью λ , то распределение моментов времени t_k имеет распределение Эрланга

$$p_{t_k}(t) = \begin{cases} \frac{\lambda e^{-\lambda t} (\lambda t)^{k-1}}{(k-1)!}, & t > 0 \\ 0, & t \leq 0. \end{cases}$$

Доказательство.

Сумма $t_k = \sum_{m=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^m}{m!}$ представляет суммарное время ожидания k -ого события; она имеет функцию распределения

$$\begin{aligned} F_{t_k}(t) &= P\{t_k < t\} = P\{X(t) \geq k\} = 1 - P\{X(t) < k\} = \\ &= 1 - \sum_{m=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^m}{m!}, \quad t > 0. \end{aligned}$$

Дифференцируя по t , находим плотность вероятности

$$\begin{aligned} F'_{t_k}(t) &= p_{t_k}(t) = \lambda e^{-\lambda t} \left(1 + \frac{\lambda t}{1!} + \dots + \frac{(\lambda t)^{k-1}}{(k-1)!} \right) - e^{-\lambda t} \times \\ &\times \left(\lambda + \frac{2\lambda t}{2!} \lambda + \dots + \frac{(k-1)(\lambda t)^{k-2}}{(k-1)!} \lambda \right) = \lambda e^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!}, \quad t > 0. \end{aligned}$$

Полученное гамма-распределение называют распределением Эрланга.

13.5. Процесс чистого рождения

Простейшим обобщением пуассоновского процесса является процесс чистого рождения, если допустить зависимость вероятности осуществления события в данный момент от числа событий, которые уже произошли. Например, воспроизведение живых организмов, когда при соответствующих условиях — изобилии пищи, отсутствии смертности, отсутствии миграции и т.д., вероятность рождения в данный момент прямо пропорциональна размеру популяции. Определим процесс чистого рождения как марковский процесс $\{X(t)\}$, $t \geq 0$, удовлетворяющий постулатам:

- 1) $P\{X(t+h) - X(t) = 1/X(t) = k\} = \lambda_k h + o_1(h)$, $h \rightarrow 0$;
- 2) $P\{X(t+h) - X(t) = 0/X(t) = k\} = 1 - \lambda_k h + o_2(h)$, $h \rightarrow 0$;
- 3) $P\{X(t+h) - X(t) < 0/X(t) = k\} = 0$, $k \geq 0$;
- 4) $X(0) = 0$ — число рождений в момент $t = 0$.

Если $h > 0$ и $n \geq 1$, то по формуле полной вероятности, марковскому свойству и постулату 3 имеем:

$$p_n(t+h) = \sum_{k=0}^{\infty} p_k(t) P\{X(t+h) = n/X(t) = k\} =$$

$$\begin{aligned}
&= \sum_{k=0}^{\infty} p_k(t) P\{X(t+h) - X(t) = n - k/X(t) = k\} = \\
&= \sum_{k=0}^{\infty} p_k(t) P\{X(t+h) - X(t) = n - k/X(t) = k\} = \\
&= \sum_{k=0}^n p_k(t) P\{X(t+h) - X(t) = n - k/X(t) = n\}.
\end{aligned}$$

При $k = 0, 1, \dots, n-2$ имеем:

$$\begin{aligned}
&P\{X(t+h) - X(t) = n - k/X(t) = k\} \leq \\
&\leq P\{X(t+h) - X(t) \geq 2/X(t) = k\} = 0_1(h) + 0_2(h)
\end{aligned}$$

или

$$\begin{aligned}
&P\{X(t+h) - X(t) = n - k/X(t) = k\} \leq \\
&\leq P\{X(t+h) - X(t) \geq 2/X(t) = k\} = 0_3(h).
\end{aligned}$$

Таким образом,

$$\begin{aligned}
p_n(t+h) &= p_n(t)[h - \lambda_n h + 0_2(h)] + p_{n-2}(t)[\lambda_{n-1} h + 0_1(h)] + \\
&\quad + \sum_{k=0}^{n-2} p_k(t) 0_3(h)
\end{aligned}$$

или

$$p_n(t+h) - p_n(t) = p_n(t)[- \lambda_n h + 0_2(h)] + p_{n-1}(t)[\lambda_{n-1} h + 0_1(h) + 0_n(h)].$$

Деля на h и переходя к пределу при $h \rightarrow 0$, получим:

$$\begin{cases} \frac{\partial p_0(t)}{\partial t} = \lambda_0 p_0(t) \\ \frac{\partial p_n(t)}{\partial t} = -\lambda_n p_n(t) + \lambda_{n-1} p_{n-1}(t) \end{cases}$$

с начальными условиями $p_0(0) = 1$, $p_n(0) = 0$, $h > 0$. Проведя последовательное интегрирование, получим:

$$p_0(t) = \lambda_0 p_0(t),$$

$$p_1(t) = \frac{\lambda_0}{\lambda_1 - \lambda_0} \left[e^{-\lambda_0 t} - e^{-\lambda_1 t} \right],$$

$$p_2(t) = \frac{\lambda_0 \lambda_1}{\lambda_1 - \lambda_0} \left[\frac{1}{\lambda_2 - \lambda_0} (e^{-\lambda_0 t} - e^{-\lambda_1 t}) + \frac{1}{\lambda_2 - \lambda_1} (e^{-\lambda_1 t} - e^{-\lambda_2 t}) \right].$$

Можно выписать и общее решение, убедившись, что $p_k(t) \geq 0$. Однако, если λ_k растут слишком быстро при росте k , может случиться, что $\sum_{k=0}^{\infty} p_k(t) < 1$.

При каких условиях происходит быстрый рост λ_k , дает следующая теорема ([18], стр. 456).

Теорема 13.5.1 (Феллера.) *Для того, чтобы при всех значениях t решения $p_k(t)$ уравнений чистого рождения удовлетворяли соотношению $\sum_{k=0}^{\infty} p_k(t) = 1$, необходимо и достаточно, чтобы $\sum_{k=0}^{\infty} \frac{1}{\lambda_k} = \infty$.*

Доказательство.

1. Рассмотрим частичную сумму ряда $\sum_{k=0}^{\infty} p_k(t)$

$$s_n(t) = p_0(t) + p_1(t) + \cdots + p_n(t).$$

Из уравнений размножения вытекает, что

$$s'_n(t) = -\lambda_n p_n(t),$$

тогда

$$1 - s_n(t) = \lambda_n \int_0^t p_n(\tau) d\tau \quad (13.5.1.)$$

Так как все члены суммы $s_n(t)$ неотрицательны, то при каждом фиксированном t сумма $s_n(t)$ с возрастанием n не убывает. Обозначив

$$\lim_{n \rightarrow \infty} (1 - s_n(t)) = \mu(t),$$

имеем

$$\lambda_n \int_0^t p_n(\tau) d\tau \geq \mu(t),$$

и, следовательно,

$$\int_0^t s_n(\tau) d\tau \geq \mu(t) \left(\frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \cdots + \frac{1}{\lambda_n} \right).$$

Так как при любом t, n имеет место неравенство $s_n(t) \leq 1$, то

$$t \geq \mu(t) \left(\frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \cdots + \frac{1}{\lambda_n} \right).$$

Если ряд $\sum_{k=0}^{\infty} \frac{1}{\lambda_k}$ расходится, то из последнего неравенства вытекает, что $\mu(t) = 0, \forall t$, следовательно,

$$\lim_{n \rightarrow \infty} (1 - s_n(t)) = 0,$$

и, следовательно,

$$s_n(t) = \sum_{k=0}^{\infty} p_k = 1.$$

2. Обратно. Из (13.5.1) ясно, что

$$\lambda_n \int_0^t p_n(\tau) d\tau \geq 1$$

и

$$\int_0^t s_n(\tau) d\tau \leq \left(\frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \cdots + \frac{1}{\lambda_n} \right).$$

В пределе при $n \rightarrow \infty$ получаем:

$$\int_0^t (1 - \mu(t)) d\tau \leq \sum_{k=0}^{\infty} \frac{1}{\lambda_k}.$$

Если $\mu(t) = 0, \forall t$, то $\int_0^t (1 - 0) d\tau = t$, так как t произвольно, то $\sum_{k=0}^{\infty} \frac{1}{\lambda_k}$ расходится.

13.6. Процесс рождения и гибели

Одно из очевидных обобщений процессов чистого рождения состоит в том, чтобы позволить процессу $\{X(t)\}$ как возрастать, так и убывать, например, из-за гибели членов популяции. Таким образом, если в момент времени t процесс находится в состоянии i , он может через некоторый случайный отрезок времени перейти в любое из соседних состояний $i + 1$ или $i - 1$. Возникающие при

этом "процессы гибели и рождения" могут рассматриваться как процессы с непрерывным временем, служащие аналогами случайных блужданий. Стохастический граф процесса представлен на рисунке 26.

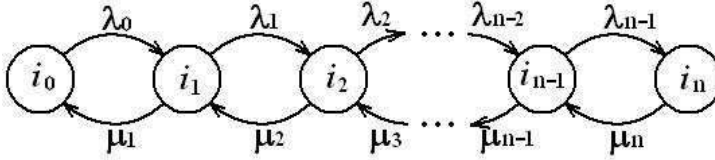


Рис. 26

Мы предположим, что $\{X(t)\}$ является марковским процессом с состояниями i_0, i_1, i_2, \dots и, что его вероятности перехода $p_{ij}(t)$ стационарны. Кроме того предположим, что $p_{ij}(t)$ удовлетворяют постулатам:

- 1) $p_{i,i+1}(h) = \lambda_i h + o(h)$, $i \geq 0$;
- 2) $p_{i,i-1}(h) = \mu_i h + o(h)$, $i \geq 1$;
- 3) $p_{i,i}(h) = 1 - (\lambda_i + \mu_i)h + o(h)$, $i \geq 0$;
- 4) $p_{ii}(0) = \delta_{ij}$;
- 5) $\mu_0 = 0$, $\lambda_0 > 0$, $\mu_i, \lambda_i > 0$, $i = 1, 2, \dots$.

Рассуждениями, подобными тем, которые были проведены выше, можно получить систему уравнений Колмогорова, управляющую процессом гибели и размножения.

$$\begin{cases} \frac{\partial p_0(t)}{\partial t} = -\lambda_0 p_0(t) + \mu_1 p_1(t) \\ \dots\dots\dots \\ \frac{\partial p_k(t)}{\partial t} = -(\lambda_k + \mu_k)p_k(t) + \mu_{k+1}p_{k+1}(t) + \lambda_{k-1}p_{k-1}(t). \end{cases} \quad (13.6.1)$$

Вопрос о существовании и единственности решения системы (13.6.1.) совсем не прост. Сформулируем свойства решений дифференциальных уравнений¹⁰.

Теорема 13.6.1. Если $\lambda_n \geq 0$ и $\mu_n \geq 0$, то существует положительное решение системы дифференциальных урав-

¹⁰Первое доказательство теоремы существования и критерия единственности было приведено в работе американского ученого Уильяма Феллера в 1940 г. Подробнее читатель может ознакомиться в учебнике [18].

Можно доказать, что при $t \rightarrow 0$ пределы существуют, не зависят от начального состояния i и удовлетворяют уравнениям Колмогорова, где правая часть равна нулю, т.е.

[illegible]

δ -распределение Дирака

Это понятие введено в двадцатых годах прошлого столетия крупнейшим английским физиком-теоретиком Полем Дираком (1902–1984), основателем квантовой механики для описания возмущений (масса, объем, заряд) в столь малом объеме, что его можно было бы принять за точку. Достаточно простые рассуждения, приводящие к этому понятию, можно найти в его книге *"Принципы квантовой механики"* [6], и суть их в следующем. Представим себе единичную массу, "размазанную" в окрестности нуля с плотностью

$$\delta_\varepsilon(x) = \begin{cases} \neq 0, & |x| \leq \frac{\varepsilon}{2} \\ = 0 & |x| > \frac{\varepsilon}{2} \end{cases}, \quad \int_{-\infty}^{\infty} \delta_\varepsilon(x) dx = 1.$$

Но тогда равенство $\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \delta_\varepsilon(x) dx = 1$ представляет собой принятый за истинный факт, что, как бы мы ни сжимали "носитель массы", сама масса должна оставаться постоянной, равной единице. А теперь попытаемся представить себе некую предельную " δ -функцию" $\delta(x)$, такую, что (см. [6])

$$\begin{cases} \int_{-\infty}^{\infty} \delta(x) dx = 1 \\ \delta(x) = 0, \quad x \neq 0. \end{cases} \quad (1)$$

Примеров функций δ_ε можно привести множество. Вот один из них:

$$\delta_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon}, & |x| \leq \frac{\varepsilon}{2} \\ 0, & |x| > \frac{\varepsilon}{2} \end{cases}, \quad \int_{-\infty}^{\infty} \delta_\varepsilon(x) dx = \frac{1}{\varepsilon} \varepsilon = 1, \quad \forall \varepsilon > 0. \quad (2)$$

Однако никак не удастся привести пример самой δ -функции. Например, если попытаться перейти к пределу при $\varepsilon \rightarrow 0$, то мы

придем к выражению $\lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(x) = \delta_0(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \neq 0 \end{cases}$, которое следует назвать противоречивым (лучше сказать "бессмысленным") по следующей причине. Исходя из физических соображений, мы должны предположить, что $\int_{-\infty}^{\infty} \delta_0(x) dx = 1$, но в математическом анализе такой интеграл известен, определен как несобственный и обязан быть равным нулю:

$$\int_{-\infty}^{\infty} \delta_0(x) dx = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{-\varepsilon} \delta_0(x) dx + \lim_{\eta \rightarrow 0} \int_{\eta}^{\infty} \delta_0(x) dx = 0,$$

поскольку на участках интегрирования под знаком предела подынтегральная функция постоянна и равна нулю. Таким образом, математические рассуждения привели к противоречию с физическим смыслом.

Ясно, что определение δ -функции (1) вызвало мощную критику (а спустя 50 лет стало ясно, что это определение привело к революции в математике, т.к. дало возможность по-другому осмыслить основы дифференциального и интегрального исчисления). П. Дирак, пытаясь уйти от критики, в [6] написал следующее: *"Дельта-функция $\delta(x)$ не является функцией в соответствии с обычным математическим определением функции, когда требуется, чтобы функция имела определенное значение для любого значения аргумента. Чтобы отметить отличие от обычного определения функции, можно назвать такую величину, как $\delta(x)$, присоединенной функцией. Таким образом, $\delta(x)$ не является такой величиной, которую можно было бы использовать повсюду в математическом анализе; ее применение должно быть ограничено некоторыми простыми типами математических выражений, для которых это применение, наверняка, не приведет к противоречиям"*.

Последней фразой П. Дирак как бы запретил использовать δ -функцию в математическом анализе(!). Сейчас мы можем сказать, что δ -функция прочно вошла в математический анализ и в его приложения (главным образом в теорию интегральных

и дифференциальных уравнений), а П. Дирак в своем определении (1) установил "физический смысл" интеграла от δ -распределения равенством

$$\int_{-\infty}^{\infty} \delta(x) dx = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \delta_{\varepsilon}(x) dx = 1.$$

Для описания взаимодействия с "точечным возмущением" (т.е. с δ -функцией, иногда говорят, "импульсом") в [6] применяется следующая схема. Сначала рассматривается лишь "размазанная плотность", а затем, после получения результата, надо перейти к пределу (если он существует), когда диаметр области, заполненной "размазанной плотностью", стремится к нулю. Такой подход приводит к результатам, которые удастся проверить на практике, и поэтому часто используется в соответствующих задачах. Надо сказать, что и с математической точки зрения эта схема дает очень разумный результат. Действительно, пусть математические преобразования с плотностью $\delta_{\varepsilon}(x)$ привели к интегральному выражению

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \delta_{\varepsilon}(x) f(x) dx$$

с непрерывной функцией $f(x)$. Применяя первую теорему о среднем ([10], стр. 459) и обозначив соответствующую среднюю точку в интервале $(-\varepsilon/2, \varepsilon/2)$ через ξ_{ε} , получим:

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \delta_{\varepsilon}(x) f(x) dx = \lim_{\varepsilon \rightarrow 0} f(\xi_{\varepsilon}) \int_{-\infty}^{\infty} \delta_{\varepsilon}(x) dx = \lim_{\varepsilon \rightarrow 0} f(\xi_{\varepsilon}) = f(0). \quad (3)$$

Эти рассуждения позволяют определить действие (неизвестной) δ -функции равенством

$$\int_{-\infty}^{\infty} f(x) \delta(x) dx = f(0). \quad (4)$$

Здесь мы видим главное, что отличает $\delta(x)$ от обычных функций. Дельта-функция Дирака — это не функция в обычном смысле,

а функционал (распределение), ставящий в соответствие каждой локально интегрируемой функции, непрерывной в окрестности нуля, ее значение в начале координат: $(\delta, \varphi) = \varphi(0)$. В современной литературе именно этот функционал и называется δ -распределением Дирака. Интегральная форма (4) равенства $(\delta, \varphi) = \varphi(0)$ должна пониматься в смысле цепочки равенств (3). (Сам Дирак в [6] трактовал равенство (4) как свойство δ -функции, которая определена равенствами (1)).

Приведенные выше рассуждения дают основание положить

$$(\delta, f) = \int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0) \quad (\delta * f) = \int_{-\infty}^{\infty} \delta(y-x) f(y) dy = f(x).$$

Дифференцирование распределений

Распределение f над классом пробных функций $\{\varphi\}$ определяется в виде $(f, \varphi) = \int_{-\infty}^{+\infty} f(x) \varphi(x) dx$. Класс пробных функций обычно подбирается по условию задачи. Наиболее распространенными являются классы $D = D(R_1)$ — множество бесконечно дифференцируемых функций с конечным носителем и $S = S(R_1)$ — множество бесконечно дифференцируемых функций убывающих вместе со всеми производными при $|x| \rightarrow \infty$ быстрее любой степени $|x|^{-1}$. Соответствующие классы распределений обозначаются D' и S' (подробности см. в учебнике [3]).

Если регулярное распределение $f \in D'(R_1)$ задается дифференцируемой функцией $f = f(x)$, интегрированием по частям, получим

$$\begin{aligned} \int_{-\infty}^{\infty} f'(x) \varphi(x) dx &= f(x) \varphi(x) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x) \varphi'(x) dx = \\ &= - \int_{-\infty}^{\infty} f(x) \varphi'(x) dx. \end{aligned}$$

Здесь выражение слева при $\varphi \in D(R_1)$ (или $\varphi \in S(R_1)$) имеет смысл для любой локально интегрируемой функции f , (порождающей медленно растущее распределение из $S'(R_1)$). Это дало

основание ввести следующее определение **производной от распределения**.

Определение. Производной от распределения $f \in D'(R_1)$ (или $f \in S'(R_1)$) называется функционал

$$f' = (f', \varphi) = -(f, \varphi'). \quad (5)$$

По индукции определяется производная порядка m :

$$(f^m, \varphi) = (-1)^m (f, \varphi^m).$$

Определение (1) позволяет находить производные распределений, задаваемых недифференцируемыми функциями. Например, пусть $f(x)$ дифференцируема везде за исключением точки x_0 , где она претерпевает разрыв первого рода. Обозначим через $[f]_{x_0}$ скачок функции $f(x)$ в точке x_0 : $[f]_{x_0} = f(x_0+0) - f(x_0-0)$. В этом случае производная от f в смысле распределений есть функционал

$$f' = \{f'(x)\} + [f]_{x_0} \delta(x-x_0), \quad (6)$$

где $\{f'\}$ — (классическая) производная от $f(t)$ в тех точках, где она дифференцируема. Действительно, интегрируя по частям, получим:

$$\begin{aligned} (f', \varphi) &= -(f, \varphi') = - \int_{-\infty}^{\infty} f(x) \varphi'(x) dx = \\ &= - \left(\int_{-\infty}^{x_0} + \int_{x_0}^{\infty} \right) f(x) \varphi'(x) dx = (f(x_0+0) - f(x_0-0)) \varphi(x_0) + \\ &+ \int_{-\infty}^{\infty} \{f'(x)\} \varphi(x) dx = ([f]_{x_0} \delta(x-x_0) + \{f'(x)\}, \varphi). \end{aligned}$$

В случае если функция имеет конечное или счетное множество точек разрыва первого рода в точках x_i и скачок функции в этих точках равен h_i , то производная соответствующего функционала определяется соотношением:

$$f' = \{f'(x)\} + \sum_i h_i \delta(x-x_i), \quad (7)$$

где $\{f'(x)\}$ производная функции $f(x)$ в точках $x \neq x_i$. В частности функция Хевисайда связана с δ – распределением следующей формулой $\Theta'(t) = \delta(t)$, где производная понимается в смысле определения (5). Аналогично $\Theta'(t - \tau) = \delta(t - \tau)$, $\forall \tau \in R_1$.

Дополнение №2. Таблицы

Значения плотности нормального распределения

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

x	p(x)	x	p(x)	x	p(x)	x	p(x)	x	p(x)	x	p(x)
0.00	0.3989	0.55	0.3429	1.10	0.2179	1.65	0.1023	2.20	0.0365	3.30	0.001723
0.01	0.3989	0.56	0.3410	1.11	0.2155	1.66	0.1006	2.22	0.0339	3.32	0.001612
0.02	0.3989	0.57	0.3391	1.12	0.2131	1.67	0.0989	2.24	0.0325	3.34	0.001508
0.03	0.3988	0.58	0.3372	1.13	0.2107	1.68	0.0973	2.26	0.0310	3.36	0.001411
0.04	0.3986	0.59	0.3352	1.14	0.2083	1.69	0.0957	2.28	0.0297	3.38	0.001319
0.05	0.3984	0.60	0.3332	1.15	0.2059	1.70	0.0940	2.30	0.0283	3.40	0.001232
0.06	0.3982	0.61	0.3312	1.16	0.2036	1.71	0.0926	2.32	0.0270	3.42	0.001151
0.07	0.3980	0.62	0.3292	1.17	0.2012	1.72	0.0909	2.34	0.0258	3.44	0.001075
0.08	0.3977	0.63	0.3271	1.18	0.1989	1.73	0.0893	2.36	0.0246	3.46	0.001003
0.09	0.3973	0.64	0.3251	1.19	0.1965	1.74	0.0879	2.38	0.0235	3.48	0.000936
0.10	0.3970	0.65	0.3230	1.20	0.1942	1.75	0.0863	2.40	0.0224	3.50	0.000873
0.11	0.3965	0.66	0.3209	1.21	0.1919	1.76	0.0848	2.42	0.0213	3.52	0.000814
0.12	0.3961	0.67	0.3187	1.22	0.1895	1.77	0.0833	2.44	0.0203	3.54	0.000759
0.13	0.3956	0.68	0.3166	1.23	0.1872	1.78	0.0818	2.46	0.0194	3.56	0.000706
0.14	0.3951	0.69	0.3144	1.24	0.1849	1.79	0.0804	2.48	0.0184	3.58	0.000657
0.15	0.3945	0.70	0.3123	1.25	0.1826	1.80	0.0790	2.50	0.0175	3.60	0.000612
0.16	0.3939	0.71	0.3101	1.26	0.1804	1.81	0.0775	2.52	0.0167	3.62	0.000569
0.17	0.3932	0.72	0.3079	1.27	0.1781	1.82	0.0761	2.54	0.0159	3.64	0.000529
0.18	0.3925	0.73	0.3056	1.28	0.1758	1.83	0.0748	2.56	0.0151	3.66	0.000492
0.19	0.3918	0.74	0.3034	1.29	0.1736	1.84	0.0734	2.58	0.0143	3.68	0.000457
0.20	0.3910	0.75	0.3011	1.30	0.1714	1.85	0.0721	2.60	0.0136	3.70	0.000425
0.21	0.3902	0.76	0.2989	1.31	0.1691	1.86	0.0707	2.62	0.0129	3.72	0.000394
0.22	0.3894	0.77	0.2966	1.32	0.1669	1.87	0.0694	2.64	0.0122	3.74	0.000365
0.23	0.3885	0.78	0.2943	1.33	0.1647	1.88	0.0681	2.66	0.0116	3.76	0.000340
0.24	0.3876	0.79	0.2920	1.34	0.1626	1.89	0.0669	2.68	0.0110	3.78	0.000315
0.25	0.3867	0.80	0.2897	1.35	0.1604	1.90	0.0656	2.70	0.0104	3.80	0.000292
0.26	0.3857	0.81	0.2874	1.36	0.1582	1.91	0.0644	2.72	0.0099	3.82	0.000271
0.27	0.3847	0.82	0.2850	1.37	0.1561	1.92	0.0632	2.74	0.0093	3.84	0.000251
0.28	0.3836	0.83	0.2827	1.38	0.1539	1.93	0.0620	2.76	0.0088	3.86	0.000232
0.29	0.3825	0.84	0.2803	1.39	0.1516	1.94	0.0608	2.78	0.0084	3.88	0.000215
0.30	0.3814	0.85	0.2780	1.40	0.1497	1.95	0.0596	2.80	0.0079	3.90	0.000199
0.31	0.3802	0.86	0.2756	1.41	0.1476	1.96	0.0584	2.82	0.0075	3.92	0.000184
0.32	0.3790	0.87	0.2732	1.42	0.1456	1.97	0.0573	2.84	0.0071	3.94	0.000170
0.33	0.3778	0.88	0.2709	1.43	0.1435	1.98	0.0562	2.86	0.0067	3.96	0.000157
0.34	0.3765	0.89	0.2685	1.44	0.1415	1.99	0.0551	2.88	0.0063	3.98	0.000145
0.35	0.3752	0.90	0.2661	1.45	0.1394	2.00	0.0540	2.90	0.0060	4.00	0.000134
0.36	0.3739	0.91	0.2637	1.46	0.1374	2.01	0.0529	2.92	0.0056	4.05	0.000109
0.37	0.3725	0.92	0.2613	1.47	0.1354	2.02	0.0518	2.94	0.0053	4.10	0.000089
0.38	0.3712	0.93	0.2589	1.48	0.1334	2.03	0.0508	2.96	0.0050	4.15	0.000073
0.39	0.3697	0.94	0.2565	1.49	0.1315	2.04	0.0498	2.98	0.0047	4.20	0.000069
0.40	0.3683	0.95	0.2541	1.50	0.1295	2.05	0.0488	3.00	0.0044	4.25	0.000048
0.41	0.3668	0.96	0.2516	1.51	0.1276	2.06	0.0478	3.02	0.0042	4.30	0.000039
0.42	0.3653	0.97	0.2492	1.52	0.1257	2.07	0.0468	3.04	0.0039	4.35	0.000031
0.43	0.3637	0.98	0.2468	1.53	0.1238	2.08	0.0459	3.06	0.0037	4.40	0.000025
0.44	0.3621	0.99	0.2444	1.54	0.1219	2.09	0.0449	3.08	0.0035	4.45	0.000020
0.45	0.3605	1.00	0.2420	1.55	0.1200	2.10	0.0440	3.10	0.0033	4.50	0.000016
0.46	0.3589	1.01	0.2396	1.56	0.1182	2.11	0.0431	3.12	0.0031	4.55	0.000013
0.47	0.3572	1.02	0.2371	1.57	0.1163	2.12	0.0422	3.14	0.0029	4.60	0.000010
0.48	0.3555	1.03	0.2347	1.58	0.1145	2.13	0.0413	3.16	0.0027	4.65	0.000008
0.49	0.3538	1.04	0.2323	1.59	0.1127	2.14	0.0404	3.18	0.0025	4.70	0.000006
0.50	0.3521	1.05	0.2299	1.60	0.1109	2.15	0.0396	3.20	0.0024	4.75	0.000005
0.51	0.3503	1.06	0.2275	1.61	0.1092	2.16	0.0387	3.22	0.0022	4.80	0.000004
0.52	0.3485	1.07	0.2251	1.62	0.1074	2.17	0.0379	3.24	0.0021	4.85	0.000003
0.53	0.3467	1.08	0.2227	1.63	0.1057	2.18	0.0371	3.26	0.0020	4.90	0.000002
0.54	0.3448	1.09	0.2203	1.64	0.1040	2.19	0.0363	3.28	0.0018	5.00	0.000001

Таб.1

Значения функции Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{y^2}{2}} dy$$

X	Φ(x)	X	Φ(x)	X	Φ(x)	X	Φ(x)	X	Φ(x)	X	Φ(x)
0,00	0,0000	0,55	0,2088	1,10	0,3643	1,65	0,4505	2,20	0,4861	3,30	0,4995
0,01	0,0040	0,56	0,2123	1,11	0,3665	1,66	0,4515	2,22	0,4868	3,32	0,4995
0,02	0,0080	0,57	0,2157	1,12	0,3686	1,67	0,4525	2,24	0,4875	3,34	0,4995
0,03	0,0120	0,58	0,2190	1,13	0,3708	1,68	0,4535	2,26	0,4881	3,36	0,4995
0,04	0,0160	0,59	0,2224	1,14	0,3729	1,69	0,4545	2,28	0,4887	3,38	0,4995
0,05	0,0199	0,60	0,2257	1,15	0,3749	1,70	0,4554	2,30	0,4893	3,40	0,4997
0,06	0,0239	0,61	0,2291	1,16	0,3770	1,71	0,4564	2,32	0,4898	3,42	0,4997
0,07	0,0279	0,62	0,2324	1,17	0,3790	1,72	0,4573	2,34	0,4904	3,44	0,4997
0,08	0,0319	0,63	0,2357	1,18	0,3810	1,73	0,4582	2,36	0,4909	3,46	0,4997
0,09	0,0359	0,64	0,2389	1,19	0,3830	1,74	0,4591	2,38	0,4913	3,48	0,4997
0,10	0,0398	0,65	0,2422	1,20	0,3849	1,75	0,4599	2,40	0,4918	3,50	0,4998
0,11	0,0438	0,66	0,2454	1,21	0,3869	1,76	0,4608	2,42	0,4922	3,52	0,4998
0,12	0,0478	0,67	0,2486	1,22	0,3888	1,77	0,4616	2,44	0,4927	3,54	0,4998
0,13	0,0517	0,68	0,2517	1,23	0,3907	1,78	0,4625	2,46	0,4931	3,56	0,4998
0,14	0,0557	0,69	0,2549	1,24	0,3925	1,79	0,4633	2,48	0,4934	3,58	0,4998
0,15	0,0596	0,70	0,2580	1,25	0,3944	1,80	0,4641	2,50	0,4938	3,60	0,4998
0,16	0,0636	0,71	0,2611	1,26	0,3962	1,81	0,4649	2,52	0,4941	3,62	0,4999
0,17	0,0675	0,72	0,2642	1,27	0,3980	1,82	0,4656	2,54	0,4945	3,64	0,4999
0,18	0,0714	0,73	0,2673	1,28	0,3997	1,83	0,4664	2,56	0,4948	3,66	0,4999
0,19	0,0753	0,74	0,2704	1,29	0,4015	1,84	0,4671	2,58	0,4951	3,68	0,4999
0,20	0,0793	0,75	0,2734	1,30	0,4032	1,85	0,4678	2,60	0,4953	3,70	0,4999
0,21	0,0832	0,76	0,2764	1,31	0,4049	1,86	0,4686	2,62	0,4956	3,72	0,4999
0,22	0,0871	0,77	0,2794	1,32	0,4066	1,87	0,4693	2,64	0,4959	3,74	0,4999
0,23	0,0910	0,78	0,2823	1,33	0,4082	1,88	0,4699	2,66	0,4961	3,76	0,4999
0,24	0,0948	0,79	0,2852	1,34	0,4099	1,89	0,4706	2,68	0,4963	3,78	0,4999
0,25	0,0987	0,80	0,2881	1,35	0,4115	1,90	0,4713	2,70	0,4965	3,80	0,4999
0,26	0,1026	0,81	0,2910	1,36	0,4131	1,91	0,4719	2,72	0,4967	3,82	0,4999
0,27	0,1064	0,82	0,2939	1,37	0,4147	1,92	0,4726	2,74	0,4969	3,84	0,4999
0,28	0,1103	0,83	0,2967	1,38	0,4162	1,93	0,4732	2,76	0,4971	3,86	0,4999
0,29	0,1141	0,84	0,2995	1,39	0,4177	1,94	0,4738	2,78	0,4973	3,88	0,4999
0,30	0,1179	0,85	0,3023	1,40	0,4192	1,95	0,4744	2,80	0,4974	3,90	0,5000
0,31	0,1217	0,86	0,3051	1,41	0,4207	1,96	0,4750	2,82	0,4976	3,92	
0,32	0,1255	0,87	0,3078	1,42	0,4222	1,97	0,4756	2,84	0,4977	3,94	
0,33	0,1293	0,88	0,3106	1,43	0,4236	1,98	0,4761	2,86	0,4979	3,96	
0,34	0,1331	0,89	0,3133	1,44	0,4251	1,99	0,4767	2,88	0,4980	3,98	
0,35	0,1368	0,90	0,3159	1,45	0,4265	2,00	0,4772	2,90	0,4981	4,00	
0,36	0,1406	0,91	0,3186	1,46	0,4279	2,01	0,4778	2,92	0,4982	4,05	
0,37	0,1443	0,92	0,3212	1,47	0,4292	2,02	0,4783	2,94	0,4984	4,10	
0,38	0,1480	0,93	0,3238	1,48	0,4306	2,03	0,4788	2,96	0,4985	4,15	
0,39	0,1517	0,94	0,3264	1,49	0,4319	2,04	0,4793	2,98	0,4986	4,20	
0,40	0,1554	0,95	0,3289	1,50	0,4332	2,05	0,4798	3,00	0,4987	4,25	
0,41	0,1591	0,96	0,3315	1,51	0,4345	2,06	0,4803	3,02	0,4987	4,30	
0,42	0,1628	0,97	0,3340	1,52	0,4357	2,07	0,4808	3,04	0,4988	4,35	
0,43	0,1664	0,98	0,3365	1,53	0,4370	2,08	0,4812	3,06	0,4989	4,40	
0,44	0,1700	0,99	0,3389	1,54	0,4382	2,09	0,4817	3,08	0,4990	4,45	
0,45	0,1736	1,00	0,3413	1,55	0,4394	2,10	0,4821	3,10	0,4990	4,50	
0,46	0,1772	1,01	0,3438	1,56	0,4406	2,11	0,4826	3,12	0,4991	4,55	
0,47	0,1808	1,02	0,3461	1,57	0,4418	2,12	0,4830	3,14	0,4992	4,60	
0,48	0,1844	1,03	0,3485	1,58	0,4429	2,13	0,4834	3,16	0,4992	4,65	
0,49	0,1879	1,04	0,3508	1,59	0,4441	2,14	0,4838	3,18	0,4993	4,70	
0,50	0,1915	1,05	0,3531	1,60	0,4452	2,15	0,4842	3,20	0,4993	4,75	
0,51	0,1950	1,06	0,3554	1,61	0,4463	2,16	0,4846	3,22	0,4994	4,80	
0,52	0,1985	1,07	0,3577	1,62	0,4474	2,17	0,4850	3,24	0,4994	4,85	
0,53	0,2019	1,08	0,3599	1,63	0,4484	2,18	0,4854	3,26	0,4994	4,90	
0,54	0,2054	1,09	0,3621	1,64	0,4495	2,19	0,4857	3,28	0,4995	5,00	

Таб. 2

Значения функции вероятности распределения Пуассона

$$p_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$$

λ	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
k										
0	0,904837	0,818731	0,740818	0,670320	0,606531	0,548812	0,496585	0,449329	0,406570	0,367879
1	0,090484	0,163746	0,222246	0,268128	0,303265	0,329287	0,347610	0,359463	0,365913	0,367879
2	0,004524	0,016375	0,033337	0,053626	0,075816	0,098786	0,121663	0,143795	0,164661	0,183940
3	0,000151	0,001092	0,003334	0,007150	0,012636	0,019757	0,028388	0,038343	0,049398	0,061313
4	0,000004	0,000055	0,000250	0,000715	0,001580	0,002964	0,004968	0,007669	0,011115	0,015328
5		0,000002	0,000015	0,000057	0,000158	0,000356	0,000696	0,001227	0,002001	0,003066
6			0,000001	0,000004	0,000013	0,000036	0,000081	0,000164	0,000300	0,000511
7					0,000001	0,000003	0,000008	0,000019	0,000039	0,000073
8							0,000001	0,000002	0,000004	0,000009
9										0,000001
λ	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0	10,0	11,0
k										
0	0,135335	0,049787	0,018316	0,006738	0,002479	0,000912	0,000335	0,000123	0,000046	0,000017
1	0,270671	0,149361	0,073263	0,033690	0,014873	0,006383	0,002684	0,001111	0,000454	0,000184
2	0,270671	0,224042	0,146525	0,084224	0,044618	0,022341	0,010735	0,004998	0,002270	0,001010
3	0,180447	0,224042	0,195367	0,140374	0,089235	0,052129	0,028626	0,014994	0,007567	0,003705
4	0,090224	0,168031	0,195367	0,175467	0,133853	0,091226	0,057252	0,033737	0,018917	0,010189
5	0,036089	0,100819	0,156293	0,175467	0,160623	0,127717	0,091604	0,060727	0,037833	0,022415
6	0,012030	0,050409	0,104196	0,146223	0,160623	0,149003	0,122138	0,091090	0,063055	0,041095
7	0,003437	0,021604	0,059540	0,104446	0,137677	0,149003	0,139587	0,117116	0,090079	0,064677
8	0,000859	0,008102	0,029770	0,065278	0,103258	0,130377	0,139587	0,131756	0,112599	0,088794
9	0,000191	0,002701	0,013231	0,036266	0,068838	0,101405	0,124077	0,131756	0,125110	0,108526
10	0,000038	0,000810	0,005292	0,018133	0,041303	0,070983	0,099262	0,118580	0,125110	0,119378
11	0,000007	0,000221	0,001925	0,008242	0,022529	0,045171	0,072190	0,097020	0,113736	0,119378
12	0,000001	0,000055	0,000642	0,003434	0,011264	0,026350	0,048127	0,072765	0,094780	0,109430
13		0,000013	0,000197	0,001321	0,005199	0,014188	0,029616	0,050376	0,072908	0,092595
14		0,000003	0,000056	0,000472	0,002228	0,007094	0,016924	0,032384	0,052077	0,072753
15		0,000001	0,000015	0,000157	0,000891	0,003311	0,009026	0,019431	0,034718	0,053352
16			0,000004	0,000049	0,000334	0,001448	0,004513	0,010930	0,021699	0,036680
17			0,000001	0,000014	0,000118	0,000596	0,002124	0,005786	0,012764	0,023734
18				0,000004	0,000039	0,000232	0,000944	0,002893	0,007091	0,014504
19				0,000001	0,000012	0,000085	0,000397	0,001370	0,003732	0,008397
20					0,000004	0,000030	0,000159	0,000617	0,001866	0,004618
21					0,000001	0,000010	0,000061	0,000264	0,000889	0,002419
22					0,000000	0,000003	0,000022	0,000108	0,000404	0,001210
23					0,000000	0,000001	0,000008	0,000042	0,000176	0,000578
24					0,000000	0,000000	0,000003	0,000016	0,000073	0,000265
25					0,000000	0,000000	0,000001	0,000006	0,000029	0,000117

Таб. 3

Значения плотности и функции показательного распределения

$$p(x) = \lambda e^{-\lambda x}, \quad x \geq 0, \quad F(x) = 1 - \lambda e^{-\lambda x}, \quad x \geq 0$$

λ	1		2		3		4		5	
x	$p(x)$	$F(x)$	$p(x)$	$F(x)$	$p(x)$	$F(x)$	$p(x)$	$F(x)$	$p(x)$	$F(x)$
0,05	0,9512	0,0488	1,8097	-0,8097	2,5821	-1,5821	3,2749	-2,2749	3,8940	-2,8940
0,10	0,9048	0,0952	1,6375	-0,6375	2,2225	-1,2225	2,6813	-1,6813	3,0327	-2,0327
0,20	0,8187	0,1813	1,3406	-0,3406	1,6464	-0,6464	1,7973	-0,7973	1,8394	-0,8394
0,30	0,7408	0,2592	1,0976	-0,0976	1,2197	-0,2197	1,2048	-0,2048	1,1157	-0,1157
0,40	0,6703	0,3297	0,8997	0,1013	0,9035	0,0965	0,8075	0,1924	0,6757	0,3233
0,50	0,6065	0,3935	0,7358	0,2642	0,5694	0,3306	0,5413	0,4587	0,4104	0,5896
0,60	0,5488	0,4512	0,6024	0,3976	0,4959	0,5041	0,3629	0,6371	0,2439	0,7511
0,70	0,4966	0,5034	0,4932	0,5068	0,3674	0,6325	0,2432	0,7568	0,1510	0,8490
0,80	0,4493	0,5507	0,4038	0,5962	0,2722	0,7278	0,1630	0,8370	0,0916	0,9084
0,90	0,4068	0,5932	0,3306	0,6694	0,2016	0,7984	0,1093	0,8907	0,0555	0,9445
1,00	0,3679	0,6321	0,2707	0,7293	0,1494	0,8506	0,0733	0,9267	0,0337	0,9663
1,10	0,3329	0,6671	0,2216	0,7784	0,1106	0,8894	0,0491	0,9509	0,0204	0,9796
1,20	0,3012	0,6988	0,1814	0,8186	0,0820	0,9180	0,0329	0,9671	0,0124	0,9876
1,30	0,2725	0,7275	0,1485	0,8515	0,0607	0,9393	0,0221	0,9779	0,0075	0,9925
1,40	0,2468	0,7532	0,1216	0,8784	0,0450	0,9550	0,0148	0,9852	0,0046	0,9954
1,50	0,2231	0,7769	0,0996	0,9004	0,0333	0,9667	0,0099	0,9901	0,0028	0,9972
1,60	0,2019	0,7981	0,0815	0,9185	0,0247	0,9753	0,0065	0,9934	0,0017	0,9983
1,70	0,1827	0,8173	0,0667	0,9333	0,0183	0,9817	0,0045	0,9965	0,0010	0,9990
1,80	0,1653	0,8347	0,0545	0,9454	0,0135	0,9865	0,0030	0,9970	0,0006	0,9994
1,90	0,1496	0,8504	0,0447	0,9553	0,0100	0,9900	0,0020	0,9980	0,0004	0,9996
2,00	0,1363	0,8647	0,0366	0,9634	0,0074	0,9926	0,0013	0,9987	0,0002	0,9998
2,10	0,1225	0,8775	0,0300	0,9700	0,0055	0,9945	0,0009	0,9991	0,0001	0,9999
2,20	0,1108	0,8892	0,0245	0,9754	0,0041	0,9959	0,0006	0,9994		
2,30	0,1003	0,8997	0,0201	0,9799	0,0030	0,9970	0,0004	0,9996		
2,40	0,0907	0,9093	0,0165	0,9835	0,0022	0,9979	0,0003	0,9997		
2,50	0,0821	0,9179	0,0135	0,9865	0,0017	0,9983	0,0002	0,9998		
2,60	0,0743	0,9257	0,0110	0,9890	0,0012	0,9988	0,0001	0,9999		
2,70	0,0672	0,9328	0,0090	0,9910	0,0009	0,9991				
2,80	0,0608	0,9392	0,0074	0,9926	0,0007	0,9993				
2,90	0,0550	0,9450	0,0061	0,9939	0,0005	0,9995				
3,00	0,0498	0,9502	0,0050	0,9950	0,0004	0,9996				
3,10	0,0450	0,9550	0,0041	0,9959	0,0003	0,9997				
3,20	0,0408	0,9592	0,0033	0,9967	0,0002	0,9998				
3,30	0,0369	0,9631	0,0027	0,9973	0,0002	0,9998				
3,40	0,0334	0,9666	0,0022	0,9978	0,0001	0,9999				
3,50	0,0302	0,9696	0,0016	0,9982						
3,60	0,0273	0,9727	0,0015	0,9985						
3,70	0,0247	0,9763	0,0012	0,9988						
3,80	0,0224	0,9796	0,0010	0,9990						
3,90	0,0202	0,9828	0,0008	0,9992						
4,00	0,0183	0,9857	0,0007	0,9993						
4,10	0,0166	0,9884	0,0005	0,9995						
4,20	0,0150	0,9909	0,0004	0,9996						
4,30	0,0136	0,9934	0,0004	0,9996						
4,40	0,0123	0,9957	0,0003	0,9997						
4,50	0,0111	0,9979	0,0002	0,9998						
4,60	0,0101	0,9999	0,0002	0,9998						
4,70	0,0091	0,9999	0,0002	0,9998						
4,80	0,0082	0,9999	0,0001	0,9999						
4,90	0,0074	0,9999								
5,00	0,0067	0,9999								
5,50	0,0041	0,9999								
6,00	0,0025	0,9999								
6,50	0,0015	0,9999								
7,00	0,0009	0,9999								
7,50	0,0006	0,9999								
8,00	0,0003	0,9999								
9,00	0,0001	0,9999								
10,00	0,0000	1,0000								

Таб. 4

Список литературы

1. Богачев Б.М., Сысоев В.В. Теория вероятностей. Воронеж, ВГТА, 2000, с. 134.
2. Борель Э. Вероятность и достоверность (перев. с франц.). М.: Наука. 1964.
3. Владимиров В.С. Уравнения математической физики. М: Наука, 1971, с. 512.
4. Гнеденко Б.В. Курс теории вероятностей: Учебник. М.: Эдиториал УРСС, 2001, с. 320.
5. Гнеденко Б.В. Очерк по теории вероятностей: М.: Эдиториал УРСС, 2001, с. 88.
6. Дирак П. Принципы квантовой механики. М.:Наука, 1979, с.480.
7. Жевержеев В.Ф., Кальницкий Л.А., Сапогов Н.А. Специальный курс высшей математики для втузов.М: Высшая школа, 1970, с. 416.
8. Ильин В.А., Куркина А.В. Высшая математика. Издательство московского университета, 2004, с. 379.
9. Краснов М.Л., Киселев А.И., Макаренко Г.И. Функции комплексного переменного. Операционное исчисление. Теория устойчивости. М: Наука, 1971, с. 254.

-
10. Кудрявцев В.С. Курс математического анализа. Т.1-2., М: Высшая школа, 1981, с. 687(1).
 11. Мостеллер Ф., Руке Р., Томас Дж. Вероятность. М.: Мир, 1969, с. 432.
 12. Никольский С.М. Курс математического анализа. Т.1-2. М.: Наука, 1973, с. 431.
 13. Нейман Ю. Вводный курс теории вероятности и математической статистики. М: Наука. 1969, с. 448.
 14. Рихтмайер Р. Принципы современной математической физики. М: Мир. 1982, с. 486.
 15. Розанов Ю.А. Лекции по теории вероятностей. М: Наука. 1962, с. 120.
 16. Севастьянов Б.А. Курс теории вероятностей и математической статистики. Москва-Ижевск: Институт компьютерных исследований, 2004, с. 272.
 17. Стоянов Й. Контрпримеры в теории вероятностей. М.: Изд-во "Факториал" , 1999, с. 288.
 18. Феллер В. Введение в теорию вероятностей и ее приложения. т.1. М: Мир, 1984, с. 493.
 19. Чистяков В.П. Курс теории вероятностей М: Наука. 1982, с. 255.
 20. Чудесенко В.Ф. Сборник заданий по специальным курсам высшей математики. М: Высшая школа, 1983, с. 112.
 21. Шварц Л. Математические методы для физических наук. М: Мир, 1965, с. 405.
 22. Шилов Г.Е. Математический анализ. Специальный курс. М.:ГИФМЛ. 1961, с. 436.