Б.В.Гнеденко

Курс ТЕОРИИ ВЕРОЯТНОСТЕЙ

Издание седьмое, исправленное

УРСС
Москва • 2001

Введение

Цель настоящей книги состоит в изложении основ *теории вероятностей* — математической науки, изучающей закономерности случайных явлений.

Возникновение теории вероятностей относится к середине XVII века и связано с именами Гюйгенса (1629–1695), Паскаля (1623–1662), Ферма (1601–1665) и Якоба Бернулли (1654–1705). В переписке Паскаля и Ферма, вызванной задачами, поставленными азартными игроками и не укладывающимися в рамки математики того времени, выкристаллизовывались постепенно такие важные понятия, как вероятность и математическое ожидание. При этом, конечно, нужно отдавать себе ясный отчет, что выдающиеся ученые, занимаясь задачами азартных игроков, предвидели и фундаментальную натурфилософскую роль науки, изучающей случайные явления. Они были убеждены в том, что на базе массовых случайных событий могут возникать четкие закономерности. И только состояние естествознания привело к тому, что азартные игры еще долго продолжали оставаться тем почти единственным конкретным материалом, на базе которого создавались понятия и методы теории вероятностей. Это обстоятельство накладывало отпечаток и на формально-математический аппарат, посредством которого решались возникавшие в теории вероятностей задачи: он сводился исключительно к элементарно арифметическим и комбинаторным методам. Последующее развитие теории вероятностей, а также широкое привлечение ее результатов и методов исследования в естествознание, и в первую очередь в физику, показали, что классические понятия и классические методы не потеряли своего значения и в настоящее время.

Серьезные требования со стороны естествознания и общественной практики (теория ошибок наблюдений, задачи теории стрельбы, проблемы статистики, в первую очередь статистики народонаселения) привели к необходимости дальнейшего развития теории вероятностей и привлечения более развитого аналитического аппарата. Особенно значительную роль в развитии аналитических методов теории вероятностей сыграли Муавр (1667–1754), Лаплас (1749–1827), Гаусс (1777–1855), Пуассон (1781–1840). С формально-аналитической стороны к этому же направлению примыкает работа творца неевклидовой геометрии Н. И. Лобачевского (1792–1856), посвященная теории ошибок при измерениях на сфере и выполненная с целью установления геометрической системы, господствующей во вселенной.

С половины XIX столетия и приблизительно до двадцатых годов нашего века развитие теории вероятностей связано в значительной мере с именами русских ученых — П. Л. Чебышева (1821–1894), А. А. Маркова (1856–1922), А. М. Ляпунова (1857–1918). Этот успех русской науки был подготовлен деятельностью В. Я. Буняковского (1804–1889), широко культивировавшего в России исследования по применению теории вероятностей к статистике, в особенности к страховому делу и демографии. Им был написан первый в России курс теории вероятностей, оказавший большое влияние на развитие интереса к этой области науки. Основное непреходящее значение работ Чебышева, Маркова и Ляпунова в области теории вероятностей состоит в том, что ими было введено в качестве объекта систематического изучения и широко использовано понятие случайной величины. С результатами Чебышева относительно закона больших чисел, с «цепями Маркова» и с предельной теоремой Ляпунова мы познакомимся в соответствующих разделах настоящей книги.

Современное развитие теории вероятностей характеризуется всеобщим подъемом интереса к ней, а также расширением круга ее практических приложений.

В этой напряженной научной работе советская школа теории вероятностей продолжает занимать выдающееся положение. Среди представителей первого поколения советских ученых прежде всего должны быть названы имена С. Н. Бернштейна (1880–1968), А. Н. Колмогорова (1903–1987) и А. Я. Хинчина (1894–1959). В процессе изложения мы будем вынуждены самим существом дела вводить читателя в курс преобразовавших лицо теории вероятностей идей и результатов. Так, уже в первой главе будем говорить о фундаментальных работах С. Н. Бернштейна, Р. Мизеса (1883–1953) и А. Н. Колмогорова по основаниям теории вероятностей. В двадцатых годах нашего столетия А. Я. Хинчин, А. Н. Колмогоров, Е. Е. Слуцкий (1880–1948) и П. Леви (1886–1971) установили тесную связь между теорией вероятностей и метрической теорией функций. Эта связь оказалась весьма плодотворной. На этом пути удалось найти окончательное решение классических задач, поставленных еще П. Л. Чебышевым, а также значительно расширить содержание теории вероятностей. Полностью к советскому периоду относится создание А. Н. Колмогоровым и А. Я. Хинчиным в тридцатых годах основ теории стохастических (вероятностных, случайных) процессов, которая теперь стала основным направлением исследований в теории вероятностей. Указанная теория служит прекрасным образцом того органического синтеза математического и естественнонаучного мышления, когда математик, овладев физическим существом узловой проблемы естествознания, находит для нее адекватный математический язык. Нам важно заметить, что решение классических задач теории вероятностей оказалось тесно связанным с теорией стохастических процессов. Элементы этой важной главы теории вероятностей будут изложены нами в главе десятой.

За последние десятилетия неизмеримо выросла роль, которую играет теория вероятностей в современном естествознании. После того как молекулярные представления о строении вещества получили всеобщее признание, стало неизбежным широкое использование теории вероятностей и в физике, и в химии. Заметим, что с точки зрения молекулярной физики каждое вещество состоит из огромного числа малых частиц, находящихся в непрерывном движении и в процессе этого движения воздействующих друг на друга. При этом о природе этих частиц, о существующем между ними взаимодействии, характере их движения и пр. известно мало. В основных чертах эти сведения исчерпываются тем, что частиц, из которых состоит вещество, очень много и что в однородном теле они близки по своим свойствам. Естественно, что при таких условиях обычные для физических теорий методы математических исследований становились бессильными. Так, например, аппарат дифференциальных уравнений не мог привести в указанной обстановке к серьезным результатам. Действительно, ни строение, ни законы взаимодействия между частицами вещества в достаточной мере не изучены, и при таких условиях применение аппарата дифференциальных уравнений должно носить элементы грубого произвола. Но даже если бы этой трудности не существовало, уже одно количество этих частиц представляет собой такую трудность в изучении их движения, которую преодолеть с помощью обычных уравнений механики нет возможности.

К тому же и методологически такой подход неудовлетворителен. Действительно, задача, которая здесь возникает, состоит не в изучении индивидуальных движений частиц, а в изучении тех закономерностей, которые возникают в совокупностях большого числа движущихся и взаимодействующих частиц. Закономерности же, возникающие вследствие участвующих в их возникновении ингредиентов, имеют свое собственное своеобразие и не сводятся к простому суммированию индивидуальных движений. Более того, эти закономерности в известных пределах оказываются не зависящими от индивидуальных особенностей участвующих в их порождении частиц. Конечно, для изучения этих новых закономерностей должны быть найдены и соответствующие новые математические методы исследования. Какие же требования должны быть в первую очередь предъявлены к этим методам? Понятно, что в первую очередь они должны учитывать то, что изучаемое явление носит массовый характер; таким образом, для этих методов наличие большого числа взаимодействующих частиц должно представлять не дополнительную трудность, а облегчать изучение возникающих закономерностей. Далее, недостаточность знаний о природе и строении частиц, а также о характере их взаимодействия не должна ограничивать эффективности их применения. Этим требованиям лучше всего удовлетворяют методы теории вероятностей.

Чтобы сказанное не было понято ошибочно, мы еще раз подчеркнем следующее обстоятельство. Говоря, что аппарат теории вероятностей лучше приспособлен для изучения молекулярных явлений, мы ни в коей мере не хотим сказать, что философские предпосылки использования теории вероятностей в естествознании лежат в «недостаточности знаний». Основной принцип состоит в том, что при изучении «*массовых*» явлений возникают *своеобразные новые закономерности*. При изучении явлений обусловленных действием большого числа молекул, учет свойств каждой молекулы не нужен. Действительно, при изучении явлений природы необходимо отвлекаться от учета *несущественных* подробностей. Рассмотрение же всех деталей, всех существующих связей, в том числе и несущественных для данного явления, приводит лишь к тому, что само явление затемняется и овладение им отодвигается ввиду такой искусственной усложненной обстановки.

Насколько удачно произведена схематизация явлений, насколько удачно выбран математический аппарат для его изучения, мы можем судить по согласию теории с опытом, с практикой. Развитие естествознания, в частности физики, показывает, что аппарат теории вероятностей оказался весьма хорошо приспособленным к изучению многочисленных явлений природы.

Указанная связь теории вероятностей с потребностями современной физики лучше всего поясняет те причины, в силу которых в последние десятилетия теория вероятностей превратилась в одну из наиболее быстро развивающихся областей математики. Новые теоретические результаты открывают новые возможности для естественнонаучного использования метода теории вероятностей. Всестороннее изучение явлений природы толкает теорию вероятностей на разыскание новых закономерностей, порождаемых случаем. Теория вероятностей не отмежевывается от запросов других наук, а идет в ногу с общим развитием естествознания. Понятно, что сказанное не означает, что теория вероятностей является лишь вспомогательным средством для решения тех или иных практических задач. Наоборот, следует подчеркнуть, что за последние три десятилетия теория вероятностей превратилась в стройную математическую дисциплину с собственными проблемами и методами доказательств. В то же время выяснилось, что наиболее существенные проблемы теории вероятностей служат делу решения различных задач естествознания.

Мы определили в самом начале теорию вероятностей как науку, изучающую случайные явления. Отложив выяснение смысла понятия «случайное явление (событие)» до первой главы, мы сейчас ограничимся несколькими замечаниями. Если в обыденных представлениях, в житейской практике считается, что случайные события представляют собой нечто крайне редкое, идущее вразрез установившемуся порядку вещей, закономерному развитию событий, то в теории вероятностей мы откажемся от этих представлений. Случайные события, как они понимаются в теории вероятностей, обладают рядом характерных особенностей; в частности, все они происходят в массовых явлениях. Под массовыми явлениями мы понимаем такие, которые имеют место в совокупностях большого числа равноправных или почти равноправных объектов и определяются именно этим массовым характером явлений и лишь в незначительной мере зависят от природы составляющих объектов.

Теория вероятностей, подобно другим разделам математики, развилась из потребностей практики: в абстрактной форме она отражает закономерности, присущие случайным событиям массового характера. Эти закономерности играют исключительно важную роль в физике и в других областях естествознания, военном деле, разнообразнейших технических дисциплинах, экономике и т. д. В последнее время в связи с широким развитием предприятий, производящих массовую продукцию, результаты теории вероятностей используются не только для браковки уже изготовленной продукции, но, что важнее, для организации самого процесса производства (статистический контроль в производстве). Большое значение в этом круге идей имеет разработка статистических методов управления качеством продукции в процессе производства. Для всего инженерного дела серьезную роль приобрела теория надежности, широко использующая методы теории вероятностей. Здесь уместно заметить, что в свою очередь теория надежности выдвинула перед теорией вероятностей ряд новых теоретических вопросов. Связь теории вероятностей с практическими потребностями, как уже было отмечено, была основной причиной бурного развития ее в последние десятилетия. Многие ее разделы были развиты как раз в связи с ответами на запросы практиков. Здесь кстати вспомнить замечательные слова основателя нашей отечественной школы теории вероятностей П. Л. Чебышева: «Сближение теории с практикой дает самые благотворные результаты, и не одна только практика от этого выигрывает; сами науки развиваются под влиянием ее: она открывает им новые предметы для исследования или новые стороны в предметах давно известных... Если теория много выигрывает от новых приложений старой методы или от новых развитий ее, то она еще более приобретает открытием новых метод, и в этом случае наука находит себе верного руководителя в практике».

Глава 1

Случайные события и их вероятности

§ 1. Интуитивные представления о случайных событиях

На протяжении длительного времени человечество изучало и использовало для своей деятельности лишь так называемые детерминистические закономерности. Подавляющая часть сведений, получаемых учащимися в школьных курсах физики, математики, химии, относится именно к этому кругу идей. Приведем примеры.

Если в основании пирамиды находится квадрат со стороной a и ее высота равна h, то объем пирамиды равен $a^2h/3$.

Если тело свободно падает на земную поверхность, то путь, пройденный им за t секунд после начала падения, равен $s = gt^2/2$.

Если химически чистую воду при атмосферном давлении 760 мм рт. ст. нагреть до 100° C, то вода начнет превращаться в пар.

При любых химических реакциях каких угодно веществ без обмена с окружающей средой общее количество вещества остается неизменным (закон сохранения вещества).

Число подобных примеров можно увеличивать неограниченно. Однако закономерности этого типа не в состоянии описать все разнообразие ситуаций, с которыми приходится сталкиваться в научной и практической деятельности. Для примера спросим себя: сколько дорожных происшествий произойдет завтра в Москве? Сколько вызовов от больных поступит в пункт скорой медицинской помощи? Сколько лет проживет только что родившийся младенец? Как много времени придется затратить на поиск неисправности в определенном сложном техническом устройстве? Все эти вопросы обладают одной особенностью — на них невозможно дать однозначного ответа, поскольку процессы, с которыми они связаны, по самому их существу лишены полной определенности. Действительно, дорожные происшествия зависят от огромного числа причин, которые невозможно предусмотреть: погода, состояние дорожного покрытия, освещенность, психологическое состояние водителей и пешеходов, взаимное расположение автомобилей на дороге и множество других. Аналогичные заключения мы можем высказать и относительно остальных предложенных нами вопросов.

Во всех подобных случаях мы говорим, что интересующее нас событие является случайным.

Поскольку случайные события врываются в нашу жизнь помимо нашего желания и постоянно окружают нас и, более того, поскольку все явления природы, согласно современным воззрениям, являются случайными, необходимо научиться их изучать и разработать для этой цели методы их изучения.

Сейчас уместно подчеркнуть, что со случайными явлениями мы вынуждены сталкиваться не время от времени, а постоянно, и зачастую именно они определяют структуру того или иного интересующего нас процесса. Так, при организации работы телефонной станции необходимо учитывать, что моменты поступления вызовов от абонентов, так же как и длительности разговора между абонентами, случайны.

В грузовой морской порт суда дальнего плавания поступают не точно по расписанию, а в моменты времени, нередко существенно отличные от запланированных. Точно так же длительность погрузо-разгрузочных работ (обработки судна) коренным образом зависит не только от погрузочных средств причала, но и от устройства трюма, характера и количества прибывших грузов, их упаковки и многих других обстоятельств. Таким образом, как при организации работы телефонной станции, так и при организации работы грузового порта, мы должны считаться с двойной случайностью — случайностью поступления требований (вызов абонентов, прибытия судов) на обслуживание и случайностью длительности их обслуживания. Мы видим, таким образом, что случайные события играют большую роль как в научных исследованиях, так и в практической деятельности. Более того, нередко при проектировании ответственных сооружений (телефонных узлов, морских портов и т. д.) мы должны опираться на эти случайные явления. Это обстоятельство привело к тому, что за последние три столетия, и в особенности за последние десятилетия, случайные явления были подвергнуты систематическому исследованию.

Прежде чем переходить к изложению основных результатов теории вероятностей, мы должны формализовать те понятия, с которыми она имеет дело. Пока же понятие «случайного» явления имеет лишь чисто описательный, интуитивный и весьма расплывчатый облик. Мы увидим, что теория вероятностей занимается изучением не любых событий, которые в житейской практике называются случайными, а только тех из них, которые обладают определенными свойствами.

Прежде всего, она ограничивается изучением лишь тех событий, которые в принципе могут быть осуществлены неограниченное число раз, притом в неизменных условиях. Приведем примеры. Игральная кость может быть подброшена в одних и тех же условиях столько раз, сколько заблагорассудится нам, исследователям. Мы можем производить неограниченное число наблюдений за числом вызовов абонентов телефонной сети, поступающих за четверть часа дневного времени. Выходные дни и ночные часы мы при этом исключим из рассмотрения, поскольку могут возникнуть опасения, что условия формирования потока вызовов в эти дни и часы будут отличны от будней.

Далее, теория вероятностей занимается лишь теми событиями, которые обладают так называемой *статистической устойчивостью* или, иначе, устойчивостью частот. Это требование к случайным событиям следует рассмотреть более подробно.

Представим, что производится последовательность испытаний, в каждом из которых может появиться, а может и не появиться некоторое событие A. Эти испытания производятся в одинаковых условиях, и результаты одних испытаний не оказывают влияния на результаты других (как говорят, испытания независимы). Обозначим через μ число появлений события A в каких-то n заранее назначенных номерах испытаний, например в n последовательных испытаниях: тогда *частота*, т. е. отношение μ/n при больших n для статистически устойчивых событий A близка к постоянной и лишь слегка изменяется то одной серии в n испытаний к другой.

Проверка статистической устойчивости представляет собой довольно сложную задачу, и сейчас мы не станем ею заниматься. Позднее же к этому вопросу мы будем возвращаться неоднократно.

Теперь подчеркнем ту мысль, что теория вероятностей не занимается изучением уникальных событий, которые не допускают повторений. Так, не имеет смысла говорить о том, какова вероятность, что данный студент сдаст экзамен по теории вероятностей на ближайшей экзаменационной сессии, поскольку здесь речь идет о единичном событии, повторить которое в тех же самых условиях нет возможности. Мы можем об этом событии высказывать лишь некоторые субъективные суждения, основанные на нашем знакомстве со знаниями этого студента. Теория же вероятностей ставит перед собой задачу изучения объективных закономерностей, которые не зависят от субъективных суждений того или иного лица. И как бы ни были интересны вопросы, касающиеся единичных, неповторимых событий, теория вероятностей к ним не имеет отношения, если только относительно них нет возможности провести длительные независимые испытания в одинаковых условиях. Так, все следующие высказывания —

6 августа 2007 г. в Термезе произойдет землетрясение силой 8 баллов;

к 2010 году будут найдены радикальные методы излечения всех форм рака;

в 2011 г. родится поэт, по таланту равный А. С. Пушкину —

носят характер уникальности, и ответ на них будет получен (положительный или отрицательный) в свое время. Пока же эти события носят неопределенный характер. И хотя они относятся к категории «может произойти, а может и не произойти», к ним понятия и методы теории вероятностей не имеют отношения.

Теория вероятностей изучает лишь такие случайные события, в отношении которых имеет смысл не только утверждение об их случайности, но и возможна объективная оценка доли случаев их появления. Эта оценка может быть выражена предложением вида:

1. Вероятность того, что при осуществлении определенного комплекса условий $\mathfrak{S}$ произойдет событие A, равна p [1)].

Закономерности такого рода называются *стохастическими* или *вероятностными*. С ними приходится иметь дело в самых разнообразных ситуациях, связанных с изучением как природных, так и общественных явлений, а также в самых разнообразных прикладных вопросах.

Для примера, пусть у нас имеется некоторое техническое устройство, скажем, электрическая лампочка. Нас интересует вопрос: проработает ли она безотказно 1 500 часов? Заранее мы не можем ответить ни утвердительно, ни отрицательно. Для этого у нас нет данных. Но мы можем и должны ответить иначе, что доля тех лампочек из большого числа находящихся в одинаковых условиях эксплуатации, которые проработают по меньшей мере 1 500 часов, равна p. Эта величина существенно зависит от того, на каком предприятии изготовлена лампочка и, безусловно, от того, в каких условиях ей приходится работать (исправность проводки и патрона, размеры колебаний напряжения и силы тока и пр.).

Конечно, каждая электрическая лампочка индивидуальна по своим качествам, но этих лампочек изготовляется очень много и испытать можно большое их число, притом изготовленных в одних и тех же производственных условиях. Мы встречаемся, таким образом, с повторением испытаний двух различных типов:

а) повторение испытаний для одного и того же объекта изучения (повторное извлечение шара из урны, содержащей несколько одинаковых шаров; извлечение наудачу карты из полной колоды и т. д.);

б) испытание многих сходных объектов. Именно по этому образцу на заводах производится испытание качества изготовленной продукции.

Формулировка детерминистических закономерностей, к которым все привыкли со школьных лет, звучит так:

2. При каждом осуществлении комплекса условий $\mathfrak{S}$ обязательно происходит событие A.

В следующем параграфе мы увидим, что вероятность случайного события измеряется числом, заключенным между 0 и 1. Единице соответствует те события,

[1)] Мы пока не даем определения понятия вероятности и полагаемся на имеющиеся у читателей интуитивные представления.

которые обязательно наступают при каждом осуществлении комплекса $\mathfrak{S}$. Такие события называются достоверными. Если же событие невозможно, то ему соответствует вероятность 0. Мы видим, что детерминистические закономерности можно рассматривать как частный случай стохастических, для которых вероятность p равна 0 или 1. Таким образом, стохастические закономерности являются более широкими, чем детерминистические, и позволяют точные, количественные методы применять и в тех случаях, когда о классическом детерминизме не может быть и речи.

Каждый исследователь, имеющий дело с применениями теории вероятностей к физике, биологии, инженерному делу, организации производства или любой другой конкретной области знаний, исходит в своей работе из убеждения, что вероятностные суждения выражают определенные объективные свойства реальных явлений. Поэтому утверждение, что при выполнении некоторого комплекса условий $\mathfrak{S}$ событие A имеет вероятность p, имеет серьезное познавательное значение. Оно указывает на наличие определенной, хотя и своеобразной, но от этого не менее объективной связи между комплексом условий $\mathfrak{S}$ и событием A. Даже только утверждение, что вероятность события A при осуществлении комплекса условий $\mathfrak{S}$ существует (хотя и неизвестна), является содержательным утверждением, нуждающимся в объективном обосновании и последующей проверке, если оно принято в качестве гипотезы. Философская задача состоит в выяснении природы этой связи. Ее решению посвящено большое число исследований, но задача до сих пор еще остается не решенной. Трудность этой проблемы привела к тому парадоксальному результату, что среди ученых встречается стремление не найти положительное ее решение, а снять ее, объявив вероятностные суждения имеющими отношение только к состоянию познающего субъекта (измеряющими степень его уверенности в наступлении события A и дающими основания для приписывания равных вероятностей исходов испытания при полном незнании). В последнее время такие субъективистские выводы нередко делаются по поводу событий, происходящих в условиях неопределенности, как теперь принято говорить.

§ 2. Поле событий. Классическое определение вероятности

Мы до сих пор не давали формализованного определения ни случайного события, ни его вероятности. Теперь приступим к этому, стремясь одновременно воспитать у читателя теоретико-вероятностную интуицию. Такая цель заставляет нас вводить определение вероятности постепенно, как бы повторяя исторический путь. Такой подход позволяет избежать формального восприятия, а отправляясь от простейших представлений, постепенно переходить к более сложным и общим.

Начнем с так называемого классического определения вероятности. При этом мы увидим, что оно в действительности является не определением, а скорее методом вычисления вероятностей при вполне определенных и сильно ограниченных условиях. Классическое определение исходит из предположения равновозможности как объективного свойства изучаемых явлений, основанного на их реальной симметрии. Понятие равновозможности (равновероятности) является первичным, не подлежащим формальному определению. Оно лишь поясняется рядом простых и доступных примеров.

При бросании на плоскость геометрически правильного куба, изготовленного из однородного материала, любая из шести граней (при подбрасывании наудачу) не имеет реальных преимуществ перед другими. Таким образом, если перенумеровать грани цифрами от 1 до 6, то при бросании куба могут произойти шесть равновероятных событий: выпадение граней 1, 2, 3, 4, 5, 6.

При подбрасывании однородного по плотности правильного двадцатигранника (икосаэдра) выпадение каждой грани в силу симметрии одинаково возможно. Мы имеем случай, когда возможны двадцать равновероятных исходов. Представим себе теперь, что при каждом испытании единственно возможны n *несовместимых* и *равновозможных исходов* $E_1, E_2, \ldots, E_n$. Каждый такой исход станем называть *элементарным событием*.

Слово «несовместимый» было введено в науку английским ученым Байесом (1702–1761) и означает, что если наступил какой-нибудь исход E_1, то ни один из остальных $n-1$ исходов в этом испытании наступить уже не мог. Так, если при бросании игральной кости выпала грань «3», то это означает, что при том же бросании не могла появиться грань «5».

Наряду с элементарными событиями рассматриваются также случайные события. Часто представляет интерес наступление при испытании не какого-то элементарного события, а одного из нескольких определенных элементарных событий. Например, при бросании игральной кости нас может интересовать появление граней с числом очков, бóльшим трех, т. е. появление какого-то из элементарных событий «4», «5», «6». Мы станем говорить в этом случае, что нас интересует случайное событие — выпадение числа очков, бóльшего трех. Вообще, если нас интересует появление какого-то из определенных элементарных событий, например, одного из событий $E_{i_1}, E_{i_2}, \ldots, E_{i_m}$, то мы станем говорить, что нас интересует наступление случайного события A, состоящего в выпадении какого-то из m только что указанных элементарных событий.

Вероятностью случайного события A называется отношение числа несовместимых и равновозможных элементарных событий, составляющих A (т. е. числа m), к числу всех возможных элементарных событий (т. е. к числу n). Вероятность случайного события A обозначается символом $\mathsf{P}(A)$. Согласно только что данному определению

$$\mathsf{P}(A) = \frac{m}{n}.$$

Например, при однократном бросании игральной кости полная группа попарно несовместимых и равновероятных событий состоит из событий

$$E_1, E_2, E_3, E_4, E_5, E_6,$$

которые состоят соответственно в выпадении 1, 2, 3, 4, 5, 6 очков. Событие, состоящее в выпадении четного числа очков, подразделяется на три частных случая, входящих в состав несовместимых и равновероятных событий. Поэтому вероятность события C равна

$$\mathsf{P}(C) = \frac{3}{6} = \frac{1}{2}.$$

Очевидно также, что в силу принятого определения

$$\mathsf{P}(E_i) = \frac{1}{6}, \quad 1 \leqslant i \leqslant 6,$$

$$\mathsf{P}(E_1 + E_2) = \frac{2}{6} = \frac{1}{3}$$

и т. д.

Рассмотрим теперь бросание двух костей. Если кости правильны, то выпадение каждой из 36 возможных комбинаций числа очков на первой и второй кости можно считать равновероятными. Так, скажем, вероятность выпадения в сумме 12 очков равна $1/36$. Выпадение в сумме 11 очков возможно двумя способами: на первой кости 5, а на другой 6, на первой кости 6, а на второй 5. Поэтому вероятность выпадения в сумме одиннадцати очков равна $2/36 = 1/18$. Читатель легко проверит, что вероятность выпадения той или иной суммы очков даются следующей таблицей:

Таблица 1

Число очков	2	3	4	5	6	7	8	9	10	11	12
Вероятность	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Подсчитаем теперь общее число возможных случайных событий, которое можно образовать из n элементарных. Очевидно, что можно образовать C_n^m событий, каждое из которых будет содержать по m каких-то элементарных событий $(1 \leqslant m \leqslant n)$. При $m = n$ случайное событие всегда происходит, т. е. оно является достоверным. Всего, таким образом, образовано $\sum\limits_{m=1}^{n} C_n^m = 2^n - 1$ событий. Добавим теперь ко всем построенным событиям еще одно, которому не соответствует ни одно элементарное событие, т. е. состоящее из пустого множества элементов. Очевидно, что оно никогда не может наступить (поскольку ему не соответствует ни одно элементарное событие). Это случайное событие носит название невозможного события. Таким образом, всех случайных событий в рассмотренном нами случае будет 2^n.

Ближайшие рассмотрения относятся не только к классическому определению вероятности, но и ко всем дальнейшим обобщениям. Будем считать фиксированным комплекс условий $\mathfrak{S}$ и станем рассматривать некоторую систему S событий $A, B, C \ldots$ [2], каждое из которых должно при каждом осуществлении комплекса $\mathfrak{S}$ *произойти* или *не произойти* [3]. Между событиями системы S могут существовать известные соотношения, с которыми мы постоянно будем иметь дело и которые поэтому прежде всего изучим.

1) Если при каждом осуществлении комплекса условий $\mathfrak{S}$, при котором происходит событие A, происходит и событие B, то мы будем говорить, что A *влечет за собой* [4] B, и обозначать это обстоятельство символом $\subset$:

$$A \subset B$$

или символом $\supset$:

$$B \supset A.$$

2) Если A влечет за собой B и в то же время B влечет за собой A, т. е. если при каждой реализации комплекса условий $\mathfrak{S}$ события A и B оба наступают или оба не наступают, то мы будем говорить, что события A и B *равносильны* и будем обозначать это обстоятельство символом $=$:

$$A = B.$$

3) Событие, состоящее в наступлении обоих событий A и B, будем называть *произведением* событий A и B и обозначать AB.

4) Событие, состоящее в наступлении хотя бы одного из событий A и B, будем называть *суммой* событий A и B и обозначать $A + B$.

---

[2] События в дальнейшем обозначаются латинскими прописными буквами $A, B, C, D, E \ldots$.

[3] Вместо «произойти» говорят также «появиться», «иметь место» или «наступить».

[4] Вместо A «влечет за собой B» говорят также «A является частным случаем B».

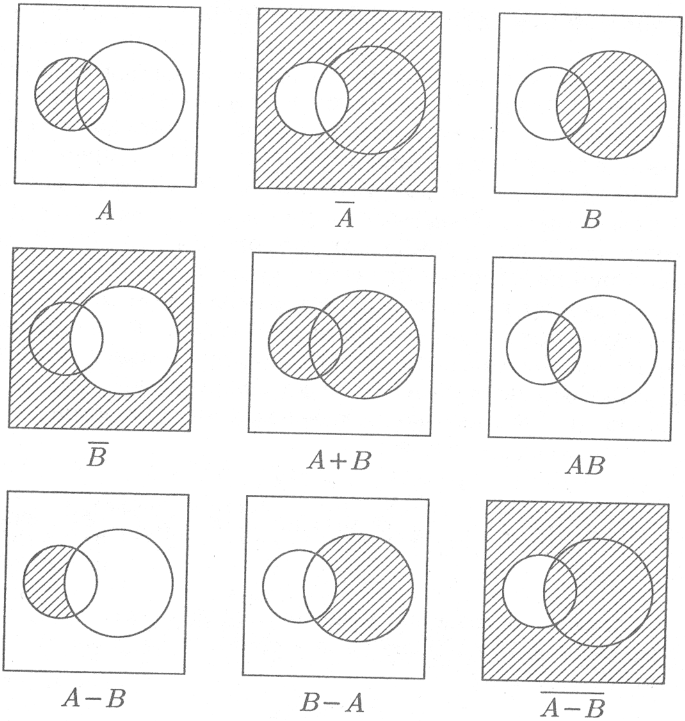

Рис. 1

5) Событие, состоящее в том, что событие A происходит, а событие B не происходит, будем называть *разностью* событий A и B и обозначать $A - B$.

6) Событие, состоящее в том, что событие A не происходит, называется *противоположным для* A и обозначается символом $\overline{A}$.

Пусть, например, комплекс условий $\mathfrak{S}$ состоит в том, что внутри квадрата, изображенного на рис. 1, выбирается наудачу точка. Обозначим через A событие «выбранная точка лежит внутри левого круга» и через B событие «выбранная точка лежит внутри правого круга». Тогда события $A, \overline{A}, B, \overline{B}, A+B, AB, A-B, B-A, \overline{A-B}$ состоят в попадании выбранной точки внутрь областей, заштрихованных на соответствующих фигурах рис. 1.

Рассмотрим другой пример. Допустим, что комплекс условий $\mathfrak{S}$ состоит в том, что на стол бросается (один раз) игральная кость. Обозначим через A выпадение на верхней грани кости[5)] шести очков, через B — выпадение трех очков, через C — выпадение какого-либо четного числа очков, через D — выпадение какого-либо числа очков, кратного трем. Тогда события A, B, C и D связаны следующими соотношениями:

$$A \subset C, \quad A \subset D, \quad B \subset D, \quad A + B = D, \quad CD = A.$$

Определение суммы и произведения двух событий обобщается на любое число событий:

$$A + B + \ldots + N$$

5) В Японии изготовляют теперь кости не только в виде кубиков, но также в виде додекаэдров и икосаэдров.

обозначает событие, заключающееся в наступлении хотя бы одного из событий $A, B, \ldots, N$, а

$$AB \ldots N$$

обозначает событие, заключающее в наступлении всех событий $A, B, \ldots, N$.

7) Событие называется *достоверным*, если оно с необходимостью должно произойти (при каждой реализации комплекса условий $\mathfrak{S}$). Например, при бросании двух игральных костей достоверно, что сумма очков будет не меньше двух.

Событие называется *невозможным*, если оно заведомо не может произойти (ни при одной реализации комплекса условий $\mathfrak{S}$). Например, при бросании двух игральных костей невозможно появление суммы очков, равной тринадцати.

Очевидно, что все достоверные события равносильны между собой. Поэтому законно обозначать все достоверные события одной буквой. Мы будем употреблять для этого букву Ω. Все невозможные события тоже равносильны между собой. Мы будем обозначать любое невозможное событие знаком $\varnothing$.

8) Два события A и $\bar{A}$ называются *противоположными*, если для них одновременно выполняются два соотношения:

$$A + \bar{A} = \Omega, \quad A\bar{A} = \varnothing .$$

Например, если при бросании одной игральной кости C обозначает выпадение четного числа очков, то

$$\Omega - C = \bar{C}$$

есть событие, состоящее в выпадении нечетного числа очков.

9) Два события A и B называются *несовместимыми*, если их совместное появление невозможно, т. е. если

$$AB = \varnothing .$$

Если

$$A = B_1 + B_2 + \ldots + B_n$$

и события B_i попарно несовместимы, т. е.

$$B_iB_j = \varnothing \quad \text{при} \quad i \neq j,$$

то говорят, что событие A *подразделяется на частные случаи* $B_1, B_2, \ldots, B_n$. Например, при бросании игральной кости событие C, состоящее в выпадении четного числа очков, подразделяется на частные случаи E_2, E_4, E_6, состоящие соответственно в выпадении 2, 4 и 6 очков.

События $B_1, B_2, \ldots, B_n$ образуют *полную группу событий*, если хотя бы одно из них непременно должно произойти (при каждом осуществлении комплекса $\mathfrak{S}$), т. е. если

$$B_1 + B_2 + \ldots + B_n = \Omega.$$

Особенно существенны для нас в дальнейшем будут *полные группы попарно несовместимых событий*. Такова, например, при однократном бросании игральной кости система событий

$$E_1, E_2, E_3, E_4, E_5, E_6,$$

состоящая, соответственно, в появлении 1, 2, 3, 4, 5 и 6 очков.

10) В каждой задаче теории вероятностей приходится иметь дело с каким-либо определенным комплексом условий $\mathfrak{S}$ и с какой-либо определенной системой S событий, наступающих или не наступающих после каждой реализации комплекса условий $\mathfrak{S}$. Относительно этой системы целесообразно сделать следующие допущения:

а) *если системе S принадлежат события A и B, то ей принадлежат также события AB, $A + B$, $A - B$*;

б) *система S содержит достоверное и невозможное события.*

Система событий, удовлетворяющая этим допущениям, называется *полем событий*.

Заметим еще в заключение, что во всех рассмотрениях теории вероятностей равносильные между собой события могут заменять друг друга. Поэтому мы условимся в дальнейшем любые два равносильных события считать просто тождественными друг другу.

Рассмотрим какую-либо полную систему G попарно несовместимых равновозможных событий (назовем их *элементарными событиями*):

$$E_1, E_2, \ldots, E_n,$$

и рассмотрим систему событий S, состоящую из невозможного события $\varnothing$, всех событий E_k системы G и всех событий A, которые могут быть подразделены на частные случаи, входящие в состав системы G.

Например, если система G состоит из трех элементарных событий E_1, E_2 и E_3, то в систему S входят события $\varnothing$, E_1, E_2, E_3, E_1+E_2, E_2+E_3, E_1+E_3, $\Omega = E_1+E_2+E_3$.

Легко установить, что система S есть поле событий. В самом деле, очевидно, что сумма, разность и произведение событий из S входят в S; невозможное событие $\varnothing$ входит в S по определению, а достоверное событие Ω входит в S так как оно представляется в виде

$$\Omega = E_1 + E_2 + \ldots + E_n.$$

В соответствии с приведенным определением (с. 19) каждому событию A, принадлежащему к построенному сейчас полю событий S, приписывается вполне определенная вероятность

$$\mathsf{P}(A) = \frac{m}{n},$$

где m есть число тех событий E_i исходной группы G, которые являются частными случаями события A. Таким образом, вероятность $\mathsf{P}(A)$ можно рассматривать как *функцию от события A, определенную на поле событий S*.

Функция эта обладает следующими свойствами:

1. *Для каждого события A поля S*

$$\mathsf{P}(A) \geqslant 0.$$

2. *Для достоверного события Ω*

$$\mathsf{P}(\Omega) = 1.$$

3. *Если событие A подразделяется на частные случаи B и C и все три события A, B и C принадлежат полю S, то*

$$\mathsf{P}(A) = \mathsf{P}(B) + \mathsf{P}(C).$$

Это свойство называется *теоремой сложения вероятностей*.

Свойство 1 очевидно, так как дробь m/n не может быть отрицательной. Свойство 2 не менее очевидно, так как достоверному событию Ω благоприятствуют все n возможных результатов испытания, и поэтому

$$\mathsf{P}(\Omega) = \frac{n}{n} = 1.$$

Докажем свойство 3. Пусть событию B благоприятствуют m', а событию C — m'' событий E_i системы G. Так как события B и C по допущению несовместимы, то события E_i, благоприятствующие одному из них, отличны от событий E_i, благоприятствующих другому. Всего, таким образом, имеется $m' + m''$ событий E_i, благоприятствующих появлению одного из событий B или C, т. е. благоприятствующих событию $B + C = A$. Следовательно,

$$\mathsf{P}(A) = \frac{m' + m''}{n} = \frac{m'}{n} + \frac{m''}{n} = \mathsf{P}(B) + \mathsf{P}(C),$$

что и требовалось доказать.

Мы ограничимся здесь указанием еще нескольких свойств вероятности.

4. *Вероятность события* $\overline{A}$, *противоположного событию* A, *равна*

$$\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A).$$

Действительно, так как

$$A + \overline{A} = \Omega,$$

то, согласно уже доказанному свойству 2,

$$\mathsf{P}(A + \overline{A}) = 1,$$

а так как события A и $\overline{A}$ несовместимы, то по свойству 3

$$\mathsf{P}(A + \overline{A}) = \mathsf{P}(A) + \mathsf{P}(\overline{A}).$$

Два последних равенства доказывают наше предположение.

5. *Вероятность невозможного события равна нулю.*

В самом деле, события Ω и $\varnothing$ несовместимы, поэтому

$$\mathsf{P}(\Omega) + \mathsf{P}(\varnothing) = \mathsf{P}(\Omega),$$

откуда следует, что

$$\mathsf{P}(\varnothing) = 0.$$

6. *Если событие* A *влечет за собой событие* B, *то*

$$\mathsf{P}(A) \leqslant \mathsf{P}(B).$$

Действительно, событие B может быть представлено как сумма двух событий A и $\overline{A}\,B$. Отсюда, в силу свойств 3 и 1, получаем:

$$\mathsf{P}(B) = \mathsf{P}(A + \overline{A}\,B) = \mathsf{P}(A) + \mathsf{P}(\overline{A}\,B) \geqslant \mathsf{P}(A).$$

7. *Вероятность любого события заключена между нулем и единицей.* Из того, что для любого события A имеют место соотношения

$$\varnothing \subset A + \varnothing = A = A\Omega \subset \Omega,$$

следует, в силу предыдущего свойства, что имеют место неравенства

$$0 = \mathsf{P}(\varnothing) \leqslant \mathsf{P}(A) \leqslant \mathsf{P}(\Omega) = 1.$$

§ 3. Примеры

Мы рассмотрим теперь несколько примеров вычисления вероятностей событий, пользуясь классическим определением вероятности. Приводимые нами примеры носят исключительно иллюстративный характер и не претендуют на то, чтобы сообщить читателю все основные методы расчета вероятностей.

Пример 1. Бросаются три игральные кости. Что вероятнее: получить в сумме выпавших очков 11 или 12?

Эта задача была одной из первичных, на которой формировались понятия и методы теории вероятностей. Утверждают, что с ней связана и следующая легенда: однажды к Галилею (а кто говорит, что к Гюйгенсу) за консультацией обратился ландскнехт. Его интересовал именно предложенный нами вопрос. Ландскнехт оказался мыслящим человеком, склонным к теоретическому и экспериментальному мышлению. Он заявил Галилею, что согласно логике обе эти суммы должны появляться одинаково часто, но опыт учит другому, а именно, что сумма 11 появляется чаще, чем 12. В чем здесь дело?

Обоснование ландскнехта на первый взгляд звучит убедительно: числа 11 и 12 оба могут быть разложены на сумму трех положительных слагаемых лишь шестью различными способами, а именно,

$$11 = 1+5+5 = 1+4+6 = 2+3+6 = 2+4+5 = 3+3+5 = 3+4+4;$$
$$12 = 1+5+6 = 2+4+6 = 2+5+5 = 3+4+5 = 3+3+6 = 4+4+4;$$

отсюда, по его мнению, вытекает равновозможность обоих интересующих нас событий.

Однако Галилей возразил ему, сказав, что каждое из этих разложений следует снабдить еще определенным весом и пояснил свою мысль таким рассуждением. Назовем кости «первой», «второй» и «третьей»; тогда разложение $1+5+5$ на самом деле может произойти не одним, а *тремя* различными способами:

$$1+5+5 = 5+1+5 = 5+5+1,$$

т. е. единица может выпасть на первой, второй или третьей кости. Точно также разложение $1+4+6$ может произойти следующими *шестью* различными способами:

$$1+4+6 = 1+6+4 = 4+1+6 = 4+6+1 = 6+1+4 = 6+4+1.$$

Таким образом, 11 в сумме может появиться не шестью, а 27 различными равновозможными способами. Сумма же 12, оказывается, разлагается лишь 25 различными способами. Здесь все дело в том, что разложение $4+4+4$ осуществимо лишь одним способом.

Теперь заметим, что общее число всех возможных равновероятных выпадений трех игральных костей равно $6^3 = 216$. Это было правильно посчитано еще в XII веке.

Обозначим через A и B случайные события, состоящие в выпадении соответственно 11 и 12 очков в сумме. Тогда согласно данному нами классическому определению вероятности,

$$\mathsf{P}(A) = \frac{27}{216} \quad \text{и} \quad \mathsf{P}(B) = \frac{25}{216}.$$

В математической статистике показывается, что для уверенного разделения двух вероятностей, отличающихся менее чем на одну сотую, нужно произвести много тысяч испытаний.

Пример 2. Из колоды карт (36 карт) наудачу вынимаются три карты. Найти вероятность того, что среди них окажется точно один туз.

Решение. Полная группа равновероятных и несовместимых событий в нашей задаче состоит из всевозможных комбинаций по три карты, их число равно C_{36}^3. Число благоприятствующих событий можно подсчитать следующим способом. Один туз мы можем выбрать C_4^1 различными способами, а две другие карты (не туз) можно выбрать C_{32}^2 различными способами. Так как для каждого определенного туза две остальные карты могут быть выбраны C_{32}^2 способами, то всего благоприятствующих случаев будет $C_4^1 \cdot C_{32}^2$. Искомая вероятность, таким образом, равна

$$p = \frac{C_4^1 \cdot C_{32}^2}{C_{36}^3} = \frac{\frac{4}{1} \cdot \frac{32 \cdot 31}{1 \cdot 2}}{\frac{36 \cdot 35 \cdot 34}{1 \cdot 2 \cdot 3}} = \frac{31 \cdot 16}{35 \cdot 3 \cdot 17} = \frac{496}{1785} \approx 0{,}2778,$$

т. е. немного больше 0,25.

Пример 3. Из колоды карт (36 карт) наудачу вынимают три карты. Найти вероятность того, что среди них окажется хотя бы один туз.

Решение. Обозначим интересующее нас событие буквой A: оно может быть представлено в виде суммы трех следующих несовместимых событий: A_1 — появление одного туза, A_2 — появление двух тузов, A_3 — появление трех тузов.

Рассуждениями, аналогичными тем, которые мы видели при решении предыдущей задачи, легко установить, что число случаев, благоприятствующих

$$\begin{array}{lll} \text{событию } A_1 & \text{равно} & C_4^1 \cdot C_{32}^2, \\ \text{-”-} \quad A_2 & \text{-”-} & C_4^2 \cdot C_{32}^1, \\ \text{-”-} \quad A_3 & \text{-”-} & C_4^3 \cdot C_{32}^0. \end{array}$$

Так как число всевозможных случаев равно C_{36}^3, то

$$\mathsf{P}(A_1) = \frac{C_4^1 \cdot C_{32}^2}{C_{36}^3} = \frac{16 \cdot 31}{3 \cdot 35 \cdot 17} \approx 0{,}2778,$$

$$\mathsf{P}(A_2) = \frac{C_4^2 \cdot C_{32}^1}{C_{36}^3} = \frac{3 \cdot 16}{3 \cdot 35 \cdot 17} \approx 0{,}0269,$$

$$\mathsf{P}(A_3) = \frac{C_4^3 \cdot C_{32}^0}{C_{36}^3} = \frac{1}{3 \cdot 35 \cdot 17} \approx 0{,}0006.$$

В силу теоремы сложения

$$\mathsf{P}(A) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) = \frac{109}{3 \cdot 119} \approx 0{,}3053.$$

Этот пример можно решить и иным методом. Событие $\overline{A}$, противоположное A, состоит в том, что среди вынутых карт не окажется ни одного туза. Очевидно, что три не-туза можно вынуть из колоды карт C_{32}^3 различными способами и, следовательно,

$$\mathsf{P}(\overline{A}) = \frac{C_{32}^3}{C_{36}^3} = \frac{32 \cdot 31 \cdot 30}{36 \cdot 35 \cdot 34} = \frac{31 \cdot 8}{3 \cdot 17 \cdot 7} \approx 0{,}6947.$$

Искомая вероятность равна

$$\mathsf{P}(A) = 1 - \mathsf{P}(\overline{A}) \approx 0{,}3053.$$

Примечание. В обоих примерах выражение «наудачу» означало, что всевозможные комбинации по три карты равновероятны.

Пример 4. Рассмотрим теперь пример, который широко используется при проверке качества принимаемой продукции. В математическом отношении он близок к примеру 2.

§ 5. О статистической оценке неизвестной вероятности

Классическое определение вероятности при переходе от простейших примеров к рассмотрению сложных задач, в особенности же задач естественнонаучного или технического характера, наталкивается на непреодолимые трудности принципиального порядка. Прежде всего, в большинстве случаев возникает вопрос о возможности нахождения разумного способа выделения «равновозможных случаев». Так, например, из соображений симметрии, на которых основаны наши суждения о равновероятности событий, вывести вероятность распада атома радиоактивного вещества за определенный промежуток времени, или же определить вероятность того, что родившийся ребенок окажется мальчиком, представляется невозможным. В этих случаях еще на заре возникновения теории вероятностей был замечен иной способ приближенной оценки неизвестной вероятности случайного события.

Длительные наблюдения над появлением или непоявлением события A при большом числе независимых испытаний, производимых при одном и том же комплексе условий $\mathfrak{S}$, в ряде случаев показывают, что число появлений события A подчиняется устойчивым закономерностям. А именно, если мы через μ обозначим число появлений события A при n независимых испытаниях, то оказывается, что отношение μ/n (частота события A) при достаточно больших n для большинства таких серий наблюдений сохраняет почти постоянную величину, причем большие отклонения наблюдаются тем реже, чем многочисленнее произведенные испытания. Более того, оказывается, что для тех случаев, к которым применимо классическое определение вероятности, это колебание частоты происходит около вероятности события p. Мы увидим впоследствии, что этот эмпирический факт имеет глубокие основания в теореме Бернулли. То, что при большом числе испытаний частота события остается почти постоянной, дает нам возможность расширить круг явлений, для которых мы будем говорить об их вероятности.

Представим себе, что относительно события A принципиально возможно проведение неограниченной последовательности не зависимых друг от друга испытаний в неизменных условиях $\mathfrak{S}$. Если в результате достаточно многочисленных наблюдений замечено, что частота события A ведет себя достаточно правильно и почти всегда колеблется около некоторой, вообще говоря неизвестной, постоянной, то мы скажем, что это событие имеет вероятность.

За численное значение этой вероятности может быть приближенно при большом числе n независимых испытаний, производящихся в неизменных условиях $\mathfrak{S}$, принята частота события A.

Однако испытания позволяют нам делать заключения и иного характера. Представим себе, что некоторые соображения дают нам основание считать, что вероятность некоторого события A равна p. Пусть далее, при проведении нескольких серий независимых испытаний оказалось, что частоты в подавляющем числе серий значительно отклоняются от величины p. Это обстоятельство дает нам основание высказать сомнение относительно правильности наших априорных суждений и предпринять более детальное исследование тех предпосылок, из которых мы исходили в своих априорных выводах. Так, например, в отношении некоторой игральной кости мы делаем предположения о ее геометрической правильности и однородности материала, из которого она изготовлена. Из этих предварительных предпосылок мы вправе сделать вывод, что при бросании кости вероятность выпадения некоторой грани, например, грани с номером 5, должна быть равна $1/6$. Если неоднократные серии достаточно многочисленных испытаний (бросаний) в нашем примере систематически показывают, что частота появления этой грани значительно отличается от $1/6$, то мы усомнимся не в существовании определенной вероятности выпадения этой грани, а в наших предпосылках о правильности кости или в правильности организации процесса наших испытаний (бросаний).

В качестве иллюстрации почти постоянства частот при больших числах испытаний рассмотрим распределение новорожденных по полу по месяцам. Данные заимствованы из книги Г. Крамера «Математические методы статистики» и представляют собой официальные данные шведской статистики за 1935 г.

Таблица 4

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	За год
Всех	7280	6957	7883	7884	7892	7609	7585	7393	7203	6903	6552	7132	88 273
Мальчиков	3743	3550	4017	4173	4117	3944	3964	3797	3712	3512	3392	3761	45 682
Девочек	3537	3407	3866	3711	3775	3665	3621	3596	3491	3391	3160	3371	42 591
Частота рождения девочек	0,486	0,489	0,490	0,471	0,478	0,482	0,462	0,484	0,485	0,491	0,482	0,473	0,4825

На рис. 9 показано уклонение частоты рождений девочек по месяцам от частоты рождений девочек за год.

Заметим, что в случае статистического определения снова имеют место такие свойства вероятности:

1) вероятность достоверного события равна единице;
2) вероятность невозможного события равна нулю;
3) если случайное событие C является суммой конечного числа несовместимых событий $A_1, A_2, \ldots, A_n$, имеющих вероятность, то его вероятность существует и равна сумме вероятностей слагаемых

$$\mathsf{P}(C) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \ldots + \mathsf{P}(A_n).$$

В заключение мы должны остановиться на весьма распространенной, в особенности среди естествоиспытателей, концепции вероятности, данной Р. Мизесом. Согласно Мизесу, раз частота по мере увеличения числа опытов все меньше и меньше уклоняется от вероятности p, то в пределе должно быть

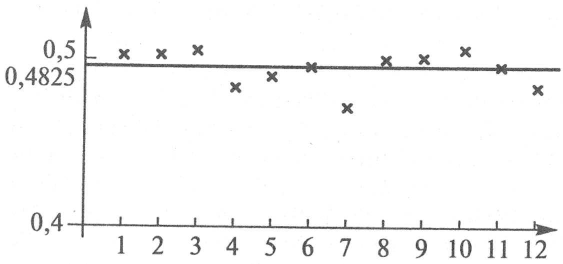

Рис. 9

$$p = \lim_{n\to\infty} \frac{\mu}{n}.$$

Это равенство Мизес предполагает считать определением понятия вероятности. По его мнению, любое априорное определение обречено на неудачу и лишь данное им эмпирическое определение способно обеспечить интересы естествознания, математики и философии, причем раз классическое определение имеет лишь весьма ограниченное применение, а статистическое определение применимо ко всем имеющим научный интерес случаям, то классическое определение через равновозможность, основанную на симметрии, Мизес предлагает вовсе отбросить. Более того, Мизес считает вообще ненужным выяснение структуры явлений, для которых вероятность является объективной числовой характеристикой, для него достаточно наличия эмпирической устойчивости частоты.

Мы не будем останавливаться на деталях теории Мизеса, в частности, на тех ограничениях и условиях, которые он дополнительно накладывает на последовательность испытаний. За подробностями теории мы отошлем читателя к его книге «Вероятность и статистика». В то же время его положения не были безоговорочно приняты наукой. Критические замечания в развернутом виде изложены в статье А. Я. Хинчина[10].

В концепции Мизеса вероятность теряет свой характер объективной числовой характеристики некоторых реальных явлений. Действительно, до производства бесконечного числа испытаний нельзя даже говорить про вероятность того или иного события, а поскольку этого нельзя осуществить, то и вообще мы лишены возможности в каких-либо условиях использовать теорию вероятностей. Следует заметить при этом, что, требуя от частот сходимости к вероятности, Мизес выставляет такое требование, какого не предъявляют ни в одной области естествознания. Ведь в самом деле, никто из нас не откажется от понятия температуры только потому, что мы не можем произвести бесконечного числа измерений и не можем проверить, будут ли результаты этих испытаний стремиться к пределу, если бы мы их все же стали производить. Или не станем же мы говорить, что какой-либо предмет не имеет размеров только потому, что последовательность наших измерений не стремиться к пределу. Более того, следуя за Мизесом, мы вообще не можем говорить о температуре тела или о существовании размеров предмета до тех пор, пока не появится мыслящий субъект, не начнет производить измерения и не убедится в том, что их результаты стремятся к пределу.

Как классическое определение вероятности, так и статистическое, были впервые четко сформулированы Я. Бернулли.

§ 6. Аксиоматическое построение теории вероятностей

Теория вероятностей долгое время представляла собой еще не сложившуюся математическую науку, в которой основные понятия были недостаточно четко определены. Эта нечеткость приводила нередко к парадоксальным выводам (вспомним парадоксы Бертрана). Естественно, что приложения теории вероятностей к изучению явлений природы были слабо обоснованы и встречали порой резкую и обоснованную критику. Нужно сказать, что эти обстоятельства мало смущали естествоиспытателей и их наивный теоретико-вероятностный подход в различных областях науки приводил к крупным успехам. Развитие естествознания в начале XX столетия предъявило к теории вероятностей повышенные требования. Возникла необходимость в систематическом изучении основных понятий теории вероятностей и выяснении тех условий, при которых возможно использование ее результатов. Вот почему особенно важное значение приобрело формально-логическое обоснование теории вероятностей, ее аксиоматическое построение. При этом в основу теории вероятностей как математической науки должны быть положены некоторые предпосылки, являющиеся обобщением многовекового человеческого опыта. Дальнейшее же ее развитие должно строиться посредством дедукции из этих основных положений без обращения к наглядным представлениям, к выводам «согласно здравому смыслу». Иными словами, теория вероятностей должна строиться из аксиом так же, как любая

[10] Частотная теория Р. Мизеса и современные идеи теории вероятности // Вопросы философии. 1961. Вып. 1. С. 91–102; Вып. 2. С. 77–89.

сформировавшаяся математическая наука — геометрия, абстрактная теория групп, теоретическая механика и т. д.

Впервые такая точка зрения была высказана и развита в 1917 г. советским математиком С. Н. Берштейном. При этом С. Н. Берштейн исходил из качественного сравнения случайных событий по их большей или меньшей вероятности.

Имеется иной подход, предложенный А. Н. Колмогоровым. Этот подход тесно связывает теорию вероятностей с современной метрической теорией функций, а также с теорией множеств. Настоящая книга следует пути, предложенному Колмогоровым.

Мы увидим, что аксиоматическое построение основ теории вероятностей отправляется от основных свойств вероятности, подмеченных на примерах классического и статистического определений. Аксиоматическое определение вероятности, таким образом, как частные случаи включает в себя и классическое и статистическое определения и преодолевает недостаточность каждого из них. На этой базе удалось построить логически совершенное здание современной теории вероятностей и в то же время удовлетворить повышенные требования к ней современного естествознания.

Отправным пунктом аксиоматики Колмогорова является множество Ω, элементы которого называются *элементарными событиями*. Наряду с Ω рассматривается множество $\mathfrak{F}$ подмножеств элементарных событий. Множество $\mathfrak{F}$ называется *алгеброй множеств*, если выполнены следующие требования:

1) $\Omega \in \mathfrak{F}$, $\varnothing \in \mathfrak{F}$ ($\varnothing$ — пустое множество);
2) из того, что $A \in \mathfrak{F}$, следует, что так же $\overline{A} \in \mathfrak{F}$;
3) из того, что $A \in \mathfrak{F}$ и $B \in \mathfrak{F}$, следует, что

$$A \cup B \in \mathfrak{F} \quad \text{и} \quad A \cap B \in \mathfrak{F}.$$

Если дополнительно к перечисленным выполняется еще следующее требование:

4) из того, что $A_n \in \mathfrak{F}$ (при $n = 1, 2, \ldots$), вытекает, что

$$\bigcup_n A_n \in \mathfrak{F} \quad \text{и} \quad \bigcap_n A_n \in \mathfrak{F},$$

то множество $\mathfrak{F}$ называется *σ-алгеброй*. Элементы $\mathfrak{F}$ называются *случайными событиями*.

Под операциями над случайными событиями понимаются операции над соответствующими множествами. В результате можно составить словарь переводов с языка теории множеств на язык теории вероятностей, приводимый нами в табл. 5.

Теперь мы можем перейти к формулировке аксиом, определяющих вероятность.

Аксиома 1. Каждому случайному событию A поставлено в соответствие неотрицательное число $\mathsf{P}(A)$, называемое его вероятностью.

Аксиома 2. $\mathsf{P}(\Omega) = 1$.

Аксиома 3 (аксиома сложения). Если события $A_1, A_2, \ldots, A_n$ попарно несовместимы, то

$$\mathsf{P}(A_1 + A_2 + \ldots + A_n) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \ldots + \mathsf{P}(A_n).$$

Для классического определения вероятности свойства, выраженные аксиомами 2 и 3, не нужно было постулировать, так как эти свойства вероятности были нами доказаны.

Из сформулированных аксиом мы выведем несколько важных элементарных следствий.

Прежде всего, из очевидного равенства

$$\Omega = \varnothing + \Omega$$

Таблица 5

Обозначения	Термины	
	теории множеств	теории вероятностей
Ω	Множество, пространство	Пространство элементарных событий, достоверное событие
ω	Элемент множества	Элементарное событие
A, B	Подмножество A, B	Случайное событие A, B
$A+B=A\cup B$	Объединение (сумма) множеств A и B	Сумма случайных событий A и B
$AB=A\cap B$	Пересечение множеств A и B	Произведение событий A и B
$\overline{A}$	Дополнение множества A	Событие, противоположное для A
$A\setminus B$	Разность множеств A и B	Разность событий A и B
$\varnothing$	Пустое множество	Невозможное событие
$AB=A\cap B=\varnothing$	Множества A и B не пересекаются (не имеют общих элементов)	События A и B несовместимы
$A=B$	Множества A и B равны	События A и B равносильны
$A\subset B$	A есть подмножество B	Событие A влечет событие B

и аксиомы 3 мы заключаем, что

$$\mathsf{P}(\Omega)=\mathsf{P}(\varnothing)+\mathsf{P}(\Omega).$$

Таким образом,

1. Вероятность невозможного события равна нулю.
2. Для любого события A

$$\mathsf{P}(\overline{A})=1-\mathsf{P}(A)\,^{11)}.$$

3. Каково бы ни было случайное событие A,

$$0\leqslant \mathsf{P}(A)\leqslant 1.$$

4. Если событие A влечет за собой событие B, то

$$\mathsf{P}(A)\leqslant \mathsf{P}(B).$$

5. Пусть A и B — два произвольных события. Поскольку в суммах $A+B=A+(B-AB)$ и $B=AB+(B-AB)$ слагаемые являются несовместимыми событиями, то в соответствии с аксиомой 3

$$\mathsf{P}(A+B)=\mathsf{P}(A)+\mathsf{P}(B-AB);\quad \mathsf{P}(B)=\mathsf{P}(AB)+\mathsf{P}(B-AB).$$

Отсюда вытекает теорема сложения для произвольных событий A и B

$$\mathsf{P}(A+B)=\mathsf{P}(A)+\mathsf{P}(B)-\mathsf{P}(AB).$$

В силу неотрицательности $\mathsf{P}(AB)$ отсюда заключаем, что

$$\mathsf{P}(A+B)\leqslant \mathsf{P}(A)+\mathsf{P}(B).$$

По индукции теперь выводим, что если $A_1, A_2, \ldots, A_n$ — произвольные события, то имеет место неравенство

$$\mathsf{P}\{A_1+A_2+\ldots+A_n\}\leqslant \mathsf{P}(A_1)+\mathsf{P}(A_2)+\ldots+\mathsf{P}(A_n).$$

[11)] Формулировка этого предложения имеется в трактате Я. Бернулли.

Система аксиом Колмогорова *непротиворечива*, так как существуют реальные объекты, которые всем этим аксиомам удовлетворяют. Например, если за Ω принять произвольное конечное множество с конечным числом элементов $\Omega = \{a_1, a_2, \dots, a_n\}$, за $\mathfrak{F}$ — совокупность всех подмножеств $\{a_{i_1}, a_{i_2}, \dots, a_{i_s}\}$, $0 \leqslant i_1 \leqslant i_2 \leqslant \dots \leqslant i_s \leqslant n$, $0 \leqslant s \leqslant n$, то положив $\mathsf{P}(a_1) = p_1, \mathsf{P}(a_2) = p_2, \dots, \mathsf{P}(a_n) = p_n$, где $p_1, p_2, \dots, p_n$ — произвольные неотрицательные числа, удовлетворяющие равенству $p_1 + p_2 + \dots + p_n = 1$, а $\mathsf{P}(a_{i_1}, a_{i_2}, \dots, a_{i_s}) = p_{i_1} + \dots + p_{i_s}$, мы удовлетворим всем аксиомам Колмогорова.

Система аксиом Колмогорова *неполна*: даже для одного и того же множества Ω вероятности в множестве $\mathfrak{F}$ мы можем выбирать различными способами.

Так, в рассмотренном нами примере с игральной костью мы можем положить или

$$\mathsf{P}(E_1) = \mathsf{P}(E_2) = \dots = \mathsf{P}(E_6) = 1/6 \tag{1}$$

или

$$\begin{gathered} \mathsf{P}(E_1) = \mathsf{P}(E_2) = \mathsf{P}(E_3) = 1/4, \\ \mathsf{P}(E_4) = \mathsf{P}(E_5) = \mathsf{P}(E_6) = 1/12, \end{gathered} \tag{2}$$

и т. д.

Неполнота системы аксиом теории вероятностей не является свидетельством их неудачного выбора или недостаточной работы мысли при их создании, а вызвана существом дела: в различных задачах могут встретиться явления, при изучении которых требуется рассматривать одинаковые множества случайных событий, но с различными вероятностями. Например, могут встретиться игральные кости, из которых одна правильная (точный куб с одинаковой плотностью в каждой точке), другая неправильная. В первом случае система вероятностей будет задана системой равенств (1), а во втором, скажем, системой (2).

Дальнейшее развитие теории нуждается в дополнительном предположении, которое носит название *расширенной аксиомы сложения*. Необходимость введения новой аксиомы объясняется тем, что в теории вероятностей постоянно приходится рассматривать события, подразделяющиеся на бесконечное число частных случаев.

Расширенная аксиома сложения. Если событие A равносильно наступлению хотя бы одного из попарно несовместимых событий $A_1, A_2, \dots, A_n, \dots$, то

$$\mathsf{P}(A) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \dots + \mathsf{P}(A_n) + \dots .$$

Заметим, что расширенная аксиома сложения может быть заменена равносильной ей *аксиомой непрерывности*.

Аксиома непрерывности. Если последовательность событий $B_1, B_2, \dots, B_n, \dots$ такова, что каждое последующее влечет за собой предыдущее и произведение всех событий B_n есть невозможное событие, то

$$\mathsf{P}(B_n) \to 0 \quad \text{при} \quad n \to \infty.$$

Докажем эквивалентность только что сформулированных предложений.

1. Из расширенной аксиомы сложения следует аксиома непрерывности. Действительно, пусть события $B_1, B_2, \dots, B_n, \dots$ таковы, что

$$B_1 \supset B_2 \supset \dots \supset B_n \supset \dots$$

и при любом $n \geqslant 1$

$$\prod_{k \geqslant n} B_k = \varnothing . \tag{3}$$

Очевидно, что

$$B_n = \sum_{k=n}^{\infty} B_k \overline{B}_{k+1} + \prod_{k=n}^{\infty} B_k.$$

Так как события, стоящие в этой сумме, попарно несовместимы, то согласно расширенной аксиоме сложения

$$\mathsf{P}(B_n) = \sum_{k=n}^{\infty} \mathsf{P}\big(B_k \overline{B}_{k+1}\big) + \mathsf{P}\left(\prod_{k=n}^{\infty} B_k\right).$$

Но в силу условия (3)

$$\mathsf{P}\left(\prod_{k=n}^{\infty} B_k\right) = 0,$$

поэтому

$$\mathsf{P}(B_n) = \sum_{k=n}^{\infty} \mathsf{P}\big(B_k \overline{B}_{k+1}\big),$$

т. е. $\mathsf{P}(B_n)$ есть остаток сходящегося ряда

$$\sum_{k=1}^{\infty} \mathsf{P}\big(B_k \overline{B}_{k+1}\big) = \mathsf{P}(B_1).$$

Поэтому $\mathsf{P}(B_n) \to 0$ при $n \to \infty$.

2. Из аксиомы непрерывности следует расширенная аксиома сложения. Пусть события $A_1, A_2, \ldots, A_n, \ldots$ попарно несовместимы и

$$A = A_1 + A_2 + \ldots + A_n + \ldots .$$

Положим

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

Ясно, что $B_{n+1} \subset B_n$. Если событие B_n наступило, то наступило какое-нибудь из событий $A_i (i \geqslant n)$ и, значит, в силу попарной несовместимости событий A_k, события $A_{i+1}, A_{i+2}, \ldots$ уже не наступили. Таким образом, события $B_{i+1}, B_{i+2}, \ldots$ невозможны, и, следовательно, невозможно событие $\prod\limits_{k=n}^{\infty} B_k$. По аксиоме непрерывности $\mathsf{P}(B_n) \to 0$ при $n \to \infty$. Так как

$$A = A_1 + A_2 + \ldots + A_n + B_{n+1},$$

то по обычной аксиоме сложения

$$\mathsf{P}(A) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \ldots + \mathsf{P}(A_n) + \mathsf{P}(B_{n+1}) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathsf{P}(A_k) = \sum_{k=1}^{\infty} \mathsf{P}(A_k).$$

Мы видим из сказанного, что аксиоматика Колмогорова позволяет строить теорию вероятностей как часть теории меры, а вероятность рассматривать как неотрицательную нормированную аддитивную функцию множества.

Вероятностным пространством принято называть тройку символов $(\Omega, \mathfrak{F}, \mathsf{P})$, где Ω — множество элементарных событий, $\mathfrak{F}$ — σ-алгебра подмножеств Ω, называемых случайными событиями, и $\mathsf{P}(A)$ — вероятность, определенная на σ-алгебре $\mathfrak{F}$.

§ 7. Условная вероятность и простейшие основные формулы

Мы уже говорили, что в основе определения вероятности события лежит некоторая совокупность условий $\mathfrak{S}$. Если никаких ограничений, кроме условий $\mathfrak{S}$, при вычислении вероятности $\mathsf{P}(A)$ не налагается, то такие вероятности называются *безусловными*.

Однако в ряде случаев приходится рассматривать вероятности событий при дополнительном условии, что произошло некоторое событие B. Такие вероятности мы будем называть *условными* и обозначать символом $\mathsf{P}(A|B)$: это означает вероятность события A при условии, что событие B произошло. Строго говоря, безусловные вероятности также являются условными, так как исходным моментом построенной теории было предположение о существовании некоторого неизменного комплекса условий $\mathfrak{S}$.

Пример 1. Брошены две игральные кости. Чему равна вероятность того, что сумма выпавших на них очков равна 8 (событие A), если известно, что эта сумма есть четное число (событие B)?

Таблица 6

(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)
(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	(6, 2)
(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)	(6, 4)
(1, 5)	(2, 5)	(3, 5)	(4, 5)	(5, 5)	(6, 5)
(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)

Все возможные случаи, которые могут представиться при бросании двух костей, мы запишем в табл. 6, каждая клетка которой содержит запись возможного события: на первом месте в скобках указывается число очков, выпавших на первой кости, на втором месте — число очков, выпавших на второй кости.

Общее число возможных случаев — 36, благоприятствующих событию A — 5. Таким образом, безусловная вероятность $\mathsf{P}(A) = 5/36$. Если событие B произошло, то осуществилась одна из 18 (а не 36) возможностей и, следовательно, условная вероятность равна $\mathsf{P}(A|B) = 5/18$.

Пример 2. Из колоды карт последовательно вынуты две карты. Найти а) безусловную вероятность того, что вторая карта окажется тузом (неизвестно, какая карта была вынута вначале) и б) условную вероятность, что вторая карта будет тузом, если первоначально был вынут туз.

Обозначим через A событие, состоящее в появлении туза на втором месте, а через B — событие, состоящее в появлении туза на первом месте. Ясно, что имеет место равенство

$$A = AB + A\overline{B}.$$

В силу несовместимости событий AB и $A\overline{B}$ имеем:

$$\mathsf{P}(A) = \mathsf{P}(AB) + \mathsf{P}\big(A\overline{B}\big).$$

При вынимании двух карт из колоды в 36 карт могут произойти $36 \cdot 35$ (учитывая порядок!) случаев. Из них благоприятствующих событию AB — $4 \cdot 3$ случаев, а событию $A\overline{B}$ — $32 \cdot 4$ случая. Таким образом,

$$\mathsf{P}(A) = \frac{4 \cdot 3}{36 \cdot 35} + \frac{32 \cdot 4}{36 \cdot 35} = \frac{1}{9}.$$

Если первая карта есть туз, то в колоде осталось 35 карт и среди них только три туза. Следовательно,

$$\mathsf{P}(A|B) = 3/35.$$

событию	A	благоприятствует	m	событий
—”—	B	—”—	k	—”—
—”—	AB	—”—	r	—”—

Общее решение задачи нахождения условной вероятности для классического определения вероятности не представляет труда. В самом деле, пусть из n единственно возможных, несовместимых и равновероятных событий $A_1, A_2, \ldots, A_n$ (понятно, что $r \leqslant k$, $r \leqslant m$). Если событие B произошло, то это означает, что наступило одно из событий A_j, благоприятствующих B. При этом условии событию A благоприятствуют r и только r событий A_j, благоприятствующих AB. Таким образом,

$$\mathsf{P}(A|B) = \frac{r}{k} = \frac{r/n}{k/n} = \frac{\mathsf{P}(AB)}{\mathsf{P}(B)}. \tag{1}$$

Точно так же, если $\mathsf{P}(A) > 0$, то

$$\mathsf{P}(B|A) = \frac{\mathsf{P}(AB)}{\mathsf{P}(A)}. \tag{1$'$}$$

Понятно, что если B (соответственно A) есть невозможное событие, то равенство (1) (соответственно (1′)) теряет смысл.

Заметим, что рассуждения, проведенные нами в примерах 1 и 2, не являются доказательствами, а представляют только мотивировки определений, данных равенствами (1) и (1′).

При $\mathsf{P}(A)\mathsf{P}(B) > 0$ каждое из равенств (1), (1′) эквивалентно так называемой теореме умножения, согласно которой

$$\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B|A) = \mathsf{P}(B)\mathsf{P}(A|B), \tag{2}$$

т. е. *вероятность произведения двух событий равна произведению вероятностей одного из этих событий на условную вероятность другого при условии, что первое произошло.*

Теорема умножения применима и в том случае, когда одно из событий A или B есть невозможное событие, так как в этом случае вместе с $\mathsf{P}(A) = 0$ имеют место равенства $\mathsf{P}(A|B) = 0$ и $\mathsf{P}(AB) = 0$.

Говорят, что событие A *независимо* от события B, если имеет место равенство

$$\mathsf{P}(A|B) = \mathsf{P}(A), \tag{3}$$

т. е. если наступление события B не изменяет вероятности события A [12)].

Если событие A независимо от B, то в силу (2) имеет место равенство

$$\mathsf{P}(A)\mathsf{P}(B|A) = \mathsf{P}(B)\mathsf{P}(A).$$

[12)] Понятия условной вероятности и независимости, а также формулировка теоремы умножения даны А. Муавром в 1718 г.

Отсюда при $\mathsf{P}(A) > 0$ находим, что

$$\mathsf{P}(B|A) = \mathsf{P}(B), \tag{4}$$

т. е. событие B также независимо от A. Таким образом, свойство независимости событий *взаимно.*

Если независимость событий A и B определить посредством равенства

$$\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B),$$

то это определение верно всегда, в том числе и тогда, когда $\mathsf{P}(A) = 0$ или $\mathsf{P}(B) = 0$.

Понятие независимости событий играет значительную роль в теории вероятностей и в ее приложениях. В частности, большая часть результатов, изложенных в настоящей книге, получена в предположении независимости тех или иных рассматриваемых событий.

В практических вопросах для определения независимости данных событий редко обращаются к проверке выполнения для них равенств (3) и (4). Обычно для этого пользуются интуитивными соображениями, основанными на опыте.

Так, например, ясно, что выпадение герба на одной монете не изменяет вероятности появления герба (решки) на другой монете, если только эти моменты во время бросания не связаны между собой (например, жестко не скреплены). Точно так же рождение мальчика у одной матери не изменяет вероятности появления мальчика (девочки) у другой матери. Эти события независимые.

Для независимых событий теорема умножения принимает особенно простой вид, а именно, *если события A и B независимы, то*

$$\mathsf{P}(AB) = \mathsf{P}(A)\cdot\mathsf{P}(B).$$

Мы обобщим теперь понятие независимости двух событий на совокупность нескольких событий.

События $B_1, B_2, \ldots, B_s$ называются *независимыми в совокупности*, если для любого события B_p из их числа и произвольных $B_{i_1}, B_{i_2}, \ldots, B_{i_r}$ из их же числа и отличных от B_p ($i_n \neq p$ и $1 \leqslant n \leqslant r$) события B_p и $B_{i_1} B_{i_2} \ldots B_{i_r}$ взаимно независимы.

В силу предыдущего, это определение эквивалентно следующему: при любых $1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant s$ и r ($1 \leqslant r \leqslant s$)

$$\mathsf{P}(B_{i_1}B_{i_2}\ldots B_{i_r}) = \mathsf{P}(B_{i_1})\mathsf{P}(B_{i_2})\ldots\mathsf{P}(B_{i_r}).$$

Заметим, что для независимости в совокупности нескольких событий недостаточно их попарной независимости. В этом можно убедиться на следующем простом примере. Представим себе, что грани тетраэдра окрашены: 1-я — в красный цвет (A), 2-я — в зеленый (B), 3-я — в синий (C) и 4-я — во все эти три цвета (ABC). Легко видеть, что вероятность грани, на которую упадет тетраэдр при бросании, в своей окраске иметь красный цвет равна $1/2$: граней четыре и две из них имеют в окраске красный цвет. Таким образом,

$$\mathsf{P}(A) = 1/2.$$

Точно так же можно посчитать, что

$$\mathsf{P}(B) = \mathsf{P}(C) = \mathsf{P}(A|B) = \mathsf{P}(B|C) = \mathsf{P}(C|A) = \mathsf{P}(B|A) = \mathsf{P}(C|B) = \mathsf{P}(A|C) = 1/2.$$

События A, B, C, таким образом, попарно независимы.

Однако если нам известно, что осуществились события B и C, то заведомо осуществилось и событие A, т. е.

$$\mathsf{P}(A|BC) = 1.$$

Таким образом, события A, B, C в совокупности зависимы.

Формула $(1')$, которая в случае классического определения была нами выведена из определения условной вероятности, в случае аксиоматического определения вероятности будет взята нами в качестве определения. Таким образом, *в общем случае* при $\mathsf{P}(A) > 0$ *по определению*

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

(В случае $\mathsf{P}(A) = 0$ условная вероятность $\mathsf{P}(B|A)$ остается неопределенной.) Это позволяет нам перенести автоматически на общее понятие вероятности все определения и результаты настоящего параграфа.

Предположим теперь, что событие B может осуществиться с одним и только с одним из n несовместимых событий $A_1, A_2, \ldots, A_n$. Иными словами, положим, что

$$B = \sum_{i=1}^{n} BA_i, \tag{5}$$

где события BA_i и BA_j с разными индексами i и j несовместимы. По теореме сложения вероятностей имеем:

$$\mathsf{P}(B) = \sum_{i=1}^{n} \mathsf{P}(BA_i).$$

Использовав теорему умножения, находим, что

$$\mathsf{P}(B) = \sum_{i=1}^{n} \mathsf{P}(A_i)\mathsf{P}(B|A_i).$$

Это равенство носит название *формулы полной вероятности* [13] и играет основную роль во всей дальнейшей теории.

В качестве иллюстрации рассмотрим два примера.

Пример 3. Имеется пять урн:

2 урны состава A_1 — по два белых шара и одному черному,
1 урна состава A_2 — по 10 черных шаров,
2 урны состава A_3 — по 3 белых шара и одному черному.

Наудачу выбирается урна и из нее наудачу вынимается шар.

Чему равна вероятность, что вынутый шар белый (событие B)?

Решение. Так как вынутый шар может быть только из урны 1-го, 2-го или 3-го состава, то

$$B = A_1B + A_2B + A_3B.$$

По формуле полной вероятности находим, что

$$\mathsf{P}(B) = \mathsf{P}(A_1)\mathsf{P}(B|A_1) + \mathsf{P}(A_2)\mathsf{P}(B|A_2) + \mathsf{P}(A_3)\mathsf{P}(B|A_3).$$

Но ясно, что

$$\mathsf{P}(A_1) = \frac{2}{5}, \qquad \mathsf{P}(A_2) = \frac{1}{5}, \qquad \mathsf{P}(A_3) = \frac{2}{5},$$

$$\mathsf{P}(B|A_1) = \frac{2}{3}, \qquad \mathsf{P}(B|A_2) = 0, \qquad \mathsf{P}(B|A_3) = \frac{3}{4}.$$

[13] Формула полной вероятности широко использовалась математиками начала XVIII века, но впервые была сформулирована как основное предложение теории вероятностей П. Лапласом лишь в конце XVIII века.

Таким образом,

$$\mathsf{P}(B) = \frac{2}{5}\cdot\frac{2}{3} + \frac{1}{5}\cdot 0 + \frac{2}{5}\cdot\frac{3}{4} = \frac{17}{30}.$$

Пример 4. Известно, что вероятность поступления k вызовов на телефонную станцию за промежуток времени длительности t равна $P_t(k)$ $(k = 0, 1, 2, \ldots)$.

Считая, что появления какого-либо числа вызовов за два соседних промежутка времени являются событиями независимыми, найти вероятность поступления s вызовов за промежуток времени длительности $2t$.

Решение. Обозначим через $A^k_{b,\,b+t}$ событие, состоящее в поступлении k вызовов за время от b до $b+t$. Очевидно, что мы имеем следующее равенство:

$$A^s_{0,\,2t} = A^0_{0,\,t}A^s_{t,\,2t} + \ldots + A^s_{0,\,t}A^0_{t,\,2t},$$

которое означает, что событие $A^s_{0,2t}$ можно рассматривать как сумму $s+1$ несовместимых событий, состоящих в том, что за первый промежуток времени длительности t поступает i вызовов, а за следующий промежуток той же продолжительности поступает $s-i$ вызовов $(i = 0, 1, 2, \ldots, s)$. По теореме сложения вероятностей

$$\mathsf{P}(A^s_{0,2t}) = \sum_{i=0}^{s} \mathsf{P}\big(A^i_{0,\,t}A^{s-i}_{t,\,2t}\big).$$

По теореме умножения вероятностей для независимых событий

$$\mathsf{P}\big(A^i_{0,\,t}A^{s-i}_{t,\,2t}\big) = \mathsf{P}\big(A^i_{0,\,t}\big)\mathsf{P}\big(A^{s-i}_{t,\,2t}\big).$$

Таким образом, если положить

$$P_{2t}(s) = \mathsf{P}\big(A^s_{0,2t}\big),$$

то

$$P_{2t}(s) = \sum_{i=0}^{s} P_t(i)\cdot P_t(s-i). \tag{6}$$

Впоследствии мы увидим, что при некоторых весьма общих условиях

$$P_t(k) = \frac{(at)^k}{k!}\exp\{-at\} \quad (k = 0, 1, 2\ldots), \tag{7}$$

где a — некоторая константа.

Из формулы (6) мы находим, что

$$P_{2t}(s) = \sum_{i=0}^{s} \frac{(at)^s \exp\{-2at\}}{i!(s-i)!} = (at)^s \exp\{-2at\} \sum_{i=0}^{s} \frac{1}{i!(s-i)!}.$$

Но

$$\sum_{i=0}^{s} \frac{1}{i!(s-i)!} = \frac{1}{s!}\sum_{i=0}^{s} \frac{s!}{i!(s-i)!} = \frac{1}{s!}(1+1)^s = \frac{2^s}{s!}.$$

Поэтому

$$P_{2t}(s) = \frac{(2at)^s \exp\{-2at\}}{s!} \quad (s = 0, 1, 2, \ldots).$$

Таким образом, если для промежутка времени длительности t имеет место формула (7), то для промежутков времени, в два раза больших, и, как легко убедиться, для любых кратных t промежутков времени характер формулы для вероятности сохраняется.

Мы в состоянии теперь вывести важные формулы Байеса или, как иногда говорят, вероятности гипотез. Пусть по-прежнему имеет место равенство (5). Требуется найти вероятность события A_i, если известно, что B произошло. Согласно теореме умножения имеем:

$$\mathsf{P}(A_iB) = \mathsf{P}(B)\mathsf{P}(A_i|B) = \mathsf{P}(A_i)\mathsf{P}(B|A_i).$$

Отсюда

$$\mathsf{P}(A_i|B) = \frac{\mathsf{P}(A_i)\mathsf{P}(B|A_i)}{\mathsf{P}(B)},$$

используя формулу полной вероятности, находим, что

$$\mathsf{P}(A_i|B) = \frac{\mathsf{P}(A_i)\mathsf{P}(B|A_i)}{\sum\limits_{j=1}^{n} \mathsf{P}(A_j)\mathsf{P}(B|A_j)}.$$

Полученные нами формулы носят название *формул Байеса*[14]. Общая схема применения этих формул к решению практических задач такова. Пусть событие B может протекать в различных условиях, относительно характера которых может быть сделано n гипотез: $A_1, A_2, \ldots, A_n$. По тем или иным причинам нам неизвестны вероятности $\mathsf{P}(A_i)$ этих гипотез до испытания. Известно также, что гипотеза A_i сообщает событию B вероятность $\mathsf{P}(B|A_i)$. Произведен опыт, в котором событие B наступило. Это должно вызвать переоценку вероятностей гипотез A_i — формулы Байеса количественно решают этот вопрос.

В артиллерийской практике производится так называемая пристрелка, имеющая своей целью уточнить наши знания относительно условий стрельбы (например, правильность прицела). В теории пристрелки широко используется формула Байеса. Мы ограничимся приведением чисто схематического примера исключительно ради иллюстрации характера задач, решаемых этой формулой.

Пример 5. Имеются пять урн следующего состава:

2 урны (состава A_1) по 2 белых и 3 черных шара,
2 урны (состава A_2) — 1 белый и 4 черных шара,
1 урна (состава A_3) — 4 белых и 1 черный шар.

Из одной наудачу выбранной урны взят шар. Он оказался белым (событие B). Чему равна после опыта вероятность (апостериорная вероятность) того, что шар вынут из урны третьего состава?

Решение. Согласно предположению

$$\mathsf{P}(A_1) = \frac{2}{5}, \qquad \mathsf{P}(A_2) = \frac{2}{5}, \qquad \mathsf{P}(A_3) = \frac{1}{5};$$
$$\mathsf{P}(B|A_1) = \frac{2}{5}, \qquad \mathsf{P}(B|A_2) = \frac{1}{5}, \qquad \mathsf{P}(B|A_3) = \frac{4}{5}.$$

Согласно формуле Байеса имеем:

$$\mathsf{P}(A_3|B) = \frac{\mathsf{P}(A_3)\mathsf{P}(B|A_3)}{\mathsf{P}(A_1)\mathsf{P}(B|A_1)+\mathsf{P}(A_2)\mathsf{P}(B|A_2)+\mathsf{P}(A_3)\mathsf{P}(B|A_3)} = \frac{\frac{1}{5}\cdot\frac{4}{5}}{\frac{2}{5}\cdot\frac{2}{5}+\frac{1}{5}\cdot\frac{2}{5}+\frac{1}{5}\cdot\frac{4}{5}} = \frac{4}{10} = \frac{2}{5}.$$

[14] Т. Байес приведенных формул не выводил, он ограничился записью формулы (1) настоящего параграфа. Приведенные формулы были выписаны лишь П. Лапласом в конце XVIII века.

Точно так же находим:

$$\mathsf{P}(A_1|B) = \frac{2}{5}, \quad \mathsf{P}(A_2|B) = \frac{1}{5}.$$

§ 8. Примеры

Мы приведем несколько более сложных примеров на использование изложенной теории.

Пример 1 [15]. Два игрока A и B продолжают некоторую игру до полного разорения одного из них. Капитал первого равняется a руб., капитал второго — b руб. Вероятность выигрыша каждой партии для игрока A равна p, а для игрока B равна q; $p + q = 1$ (ничьи отсутствуют). В каждой партии выигрыш одного игрока (и, значит, проигрыш другого) равняется 1 рублю. Найти вероятность разорения каждого из игроков (результаты отдельных партий предполагаются независимыми).

Решение. Обозначим через p_n вероятность разорения игрока A, когда он имеет n руб. Очевидно, что искомая вероятность есть p_a и что

$$p_{a+b} = 0, \quad p_0 = 1, \tag{1}$$

поскольку в первом случае игрок A уже сосредоточил в своих руках весь капитал, а во втором он уже ничего не имеет.

Если игрок A имел n руб. перед некоторой партией, то его разорение может осуществиться двумя различными способами: или он очередную партию выиграет, а всю игру проиграет, или он проиграет и партию и игру. По формуле полной вероятности поэтому

$$p_n = p \cdot p_{n+1} + q \cdot p_{n-1}.$$

Относительно p_n мы получили уравнение в конечных разностях; легко видеть, что его можно записать в следующем виде:

$$q(p_n - p_{n-1}) = p(p_{n+1} - p_n). \tag{2}$$

Рассмотрим сначала решение этого уравнения при $p = q = 1/2$. При этом допущении

$$p_{n+1} - p_n = p_n - p_{n-1} = \ldots = p_1 - p_0 = c,$$

где c — постоянная. Отсюда находим, что

$$p_n = p_0 + nc.$$

Поскольку $p_0 = 1$ и $p_{a+b} = 0$, то

$$p_n = 1 - \frac{n}{a+b}.$$

Таким образом, вероятность разорения игрока A равняется

$$p_a = 1 - \frac{a}{a+b} = \frac{b}{a+b}.$$

[15] Мы сохраняем для этой задачи «о разорении игрока» ее классическую формулировку, но возможны и иные формулировки, например: материальная частица находится на прямой в точке O и каждую секунду подвергается случайному толчку, в результате которого передвигается на 1 см вправо с вероятностью p или на 1 см влево с вероятностью $q = 1 - p$. Чему равна вероятность того, что материальная частица окажется правее точки с координатой b $(b > 0)$, прежде чем она попадет в положение, расположенное левее точки с координатой a $(a < 0$, a и b — целые числа)?

Задача о разорении игрока была предложена и впервые изучена Х. Гюйгенсом. Мы предполагаем, что вероятность события «разорение игрока» существует.

Решение этого уравнения дает нам:

$$\pi(t) = \exp\{-a\tau\}\left(c_1 + at + \frac{a^2(t-\tau)^2}{2!}\right).$$

Постоянное c_1 может быть найдено из того, что согласно (6)

$$\pi(\tau) = \exp\{-a\tau\}(1 + a\tau).$$

Таким образом, $c_1 = 1$ и для $\tau \leqslant t \leqslant 2\tau$

$$\pi(t) = \exp\{-at\}\left[1 + at + \frac{a^2(t-\tau)^2}{2!}\right].$$

Методом полной индукции можно доказать, что для $(n-1)\tau \leqslant t \leqslant n\tau$ имеет место равенство

$$\pi(t) = \exp\{-at\}\sum_{k=0}^{n}\frac{a^k[t-(k-1)\tau]^k}{k!}.$$

Упражнения

A, B, C — случайные события.

1. Каков смысл равенств

а) $ABC = A$;
б) $A + B + C = A$?

2. Упростить выражения

а) $(A+B)(B+C)$;
б) $(A+B)(A+\overline{B})$;
в) $(A+B)(A+\overline{B})(\overline{A}+B)$.

3. Доказать равенства

а) $\overline{\overline{A}\,\overline{B}} = A + B$;
б) $\overline{\overline{A}+\overline{B}} = AB$;
в) $\overline{A_1 + A_2 + \ldots + A_n} = \overline{A}_1\overline{A}_2\ldots\overline{A}_n$;
г) $\overline{A_1A_2\ldots A_n} = \overline{A}_1 + \overline{A}_2 + \ldots + \overline{A}_n$.

4. Четырехтомное сочинение расположено на полке в случайном порядке. Чему равна вероятность того, что тома стоят в должном порядке справа налево или слева направо?

5. Числа 1, 2, 3, 4, 5 написаны на пяти карточках. Наудачу последовательно вынимаются три карточки, и вынутые таким образом цифры ставятся слева направо. Чему равна вероятность того, что полученное таким образом трехзначное число окажется четным?

6. В партии, состоящей из N изделий, имеются M бракованных. Наудачу выбираются n изделий из этой партии $(n < N)$. Чему равна вероятность того, что среди них окажутся m бракованных $(m \leqslant M)$?

7. Технический контроль проверяет изделия в партии, состоящей из m изделий первого сорта и n изделий второго сорта. Проверка первых b изделий, выбранных из партии наудачу, показала, что все они второго сорта $(b < n)$. Чему равна вероятность того, что среди следующих двух наудачу выбранных из числа непроверенных изделий по меньшей мере одно окажется также второго сорта?

8. Пользуясь теоретико-вероятностными соображениями, доказать тождество

$$1 + \frac{N-n}{N-1} + \frac{(N-n)(N-n-1)}{(N-1)(N-2)} + \ldots + \frac{(N-n)\ldots 2\cdot 1}{(N-1)\ldots(n+1)n} = \frac{N}{n}.$$

Указание. Из урны, содержащей N шаров и среди них n белых, наудачу вынимаются шары без возвращения. Найти вероятность того, что рано или поздно натолкнутся на белый шар.

9. Из ящика, содержащего m белых и n черных шаров $(m > n)$, вынимают наудачу один шар за другим. Чему равна вероятность того, что наступит момент, когда число вынутых черных шаров будет равно числу вынутых белых?

10. Некто написал n адресатам письма, в каждый конверт вложил по одному письму и затем наудачу написал на каждом конверте один из n адресов. Чему равна вероятность того, что хотя бы одно письмо попало по назначению?

11. В урне имеется n билетов с номерами от 1 до n. Билеты вынимаются наудачу по одному (без возращения). Чему равна вероятность того, что хотя бы при одном вынимании номер вынутого билета совпадает с номером произведенного испытания?

12. Из урны, содержащей n белых и n черных шаров, наудачу вынимается четное число шаров (все различные способы вынуть четное число шаров, независимо от их числа, считаются равновероятными). Найти вероятность того, что среди вынутых шаров окажется одинаковое число черных и белых.

13. *Задача кавалера де Мере.* Что вероятнее: при бросании четырех игральных костей хотя бы на одной получить 1 или при 24 бросаниях двух костей хотя бы раз получить две единицы?

14. На отрезок $(0, a)$ наудачу брошены три точки. Найти вероятность того, что из отрезков, равных расстояниям от точки 0 до точек падения, можно составить треугольник.

15. Стержень длины l разломан в двух наудачу выбранных точках. Чему равна вероятность того, что из полученных отрезков можно составить треугольник?

16. На отрезок AB длины a наудачу бросается точка. На отрезок BC длины b также наудачу бросается точка. Чему равна вероятность того, что из отрезков: 1) от точки A до первой брошенной точки, 2) между двумя брошенными точками, 3) от второй брошенной точки до точки C можно составить треугольник?

17. В сфере радиуса R случайно и независимо друг от друга разбросано N точек.

а) Чему равна вероятность того, что расстояние от центра до ближайшей точки будет не менее r?

б) К чему стремится вероятность, найденная в вопросе а), если $R \to \infty$ и $\dfrac{N}{R^3} \to \dfrac{4}{3}\pi\lambda$?

Примечание. Задача заимствована из звездной астрономии: в окрестности Солнца $\lambda \approx 0{,}0063$, если R измерено в парсеках.

18. События $A_1, A_2, \ldots, A_n$ независимы; $\mathsf{P}(A_k) = p_k$. Найти вероятность

а) появления хотя бы одного из этих событий,
б) непоявления всех этих событий,
в) появления точно одного (безразлично какого) события.

19. Доказать, что если события A и B несовместимы, $\mathsf{P}(A) > 0$ и $\mathsf{P}(B) > 0$, то события A и B зависимы.

20. Пусть $A_1, A_2, \ldots, A_n$ — случайные события. Доказать формулу

$$\mathsf{P}\left\{\sum_{i=1}^{n} A_i\right\} = \sum_{i=1}^{n} \mathsf{P}(A_i) - \sum_{1\leqslant i<j\leqslant n} \mathsf{P}(A_iA_j) + \sum_{1\leqslant i<j<k\leqslant n} \mathsf{P}(A_iA_jA_k) - \ldots + (-1)^{n-1}\mathsf{P}(A_1A_2\ldots A_n).$$

Посредством этой формулы решить задачи 10 и 11.

21. Вероятность того, что молекула, испытавшая в момент $t = 0$ столкновение с другой молекулой и не имевшая других столкновений до момента t, испытывает столкновение в промежуток времени между t и $t + \Delta t$, равна $\lambda\Delta t + o(\Delta t)$. Найти вероятность того, что время свободного пробега (т. е. время между двумя соседними столкновениями) будет больше t.

22. Считая, что при размножении бактерий делением (на две бактерии) вероятность бактерии разделиться за промежуток времени Δt равна $a\Delta t + o(\Delta t)$ и не зависит от числа предшествующих делений, а также от числа имеющихся бактерий, найти вероятность того, что если в момент 0 была 1 бактерия, то в момент t окажется i бактерий.