Б. В. ГНЕДЕНКО, А. Я. ХИНЧИН

ЭЛЕМЕНТАРНОЕ ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

Б. В. ГНЕДЕНКО, А. Я. ХИНЧИН

ЭЛЕМЕНТАРНОЕ ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

ИЗДАНИЕ СЕДЬМОЕ, ДОПОЛНЕННОЕ

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1970

517.8
Г 56
УДК 519.2

Элементарное введение в теорию вероятностей. Гнеденко Б. В., Хинчин А. Я. Главная редакция физико-математической литературы изд-ва «Наука», 1970.

Настоящая книжка двух советских математиков выдержала несколько изданий в нашей стране и переведена во многих странах: Франции, ГДР, США, Польше, Венгрии, Чехословакии, Румынии, Аргентине, Японии, Испании, КНР. Повсюду она встретила благожелательное отношение читателей. Эта книжка предъявляет минимальные требования к математическим знаниям читателей. Математического образования в объеме средней школы вполне достаточно для свободного понимания всех ее разделов. Изложение ведется на базе рассмотрения примеров практического содержания. При этом, однако, авторы не стремятся углубиться в детали специально технические, чтобы не затемнять суть рассматриваемых теоретико-вероятностных вопросов.

Седьмое издание отличается от шестого исправлением замеченных опечаток и добавлением новой главы, посвященной изложению элементов теории случайных процессов, получившей уже право называться одним из основных математических орудий современной практики.

Рисунков — 18.

2-2-3
49-70

ОГЛАВЛЕНИЕ

ЧАСТЬ ПЕРВАЯ

ВЕРОЯТНОСТИ

ЧАСТЬ ВТОРАЯ
СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

ЧАСТЬ ТРЕТЬЯ
СЛУЧАЙНЫЕ ПРОЦЕССЫ

ПРЕДИСЛОВИЕ К СЕДЬМОМУ ИЗДАНИЮ

Вторично без моего учителя и соавтора я вношу изменение, написав новую главу.

Когда мы задумали написать элементарную книжку по теории вероятностей, перед нашими глазами были молодые люди, окончившие среднюю школу и отброшенные вихрем Великой Отечественной войны от общения с наукой. Позднее выяснилось, что круг читателей этой книжки оказался несравненно более широким и именно по ней знакомились с идеями и методами теории вероятностей инженеры и экономисты, биологи и лингвисты, медики и военные. Меня радует, что интерес к этой книжке не пропал как в нашей стране, так и за ее пределами. Само собой разумеется, что изменение круга читателей должно оказать некоторое влияние и на содержание книги. Поскольку для многочисленных применений теории вероятностей и для развития ее теории особую роль теперь играет теория случайных процессов, я счел необходимым дополнить книжку небольшим введением в эту важную область идей и исследований. Понятно, что, сообразуясь с общим назначением книжки, в этом дополнении обращено основное внимание не на проблемы теории или аналитический аппарат, а на общее ознакомление с реальными вопросами, приводящими к теории случайных процессов.

С большой благодарностью я приму от читателей любые пожелания, относящиеся к содержанию книжки, стилю изложения и характеру рассмотренных примеров.

Б. В. Гнеденко

Москва,
10 декабря 1969 г.

ПРЕДИСЛОВИЕ К ПЯТОМУ ИЗДАНИЮ

Настоящее издание было подготовлено мной к печати уже после смерти А. Я. Хинчина — выдающегося ученого и педагога. Современное развитие теории вероятностей, многие ее идеи и результаты тесно связаны с именем Хинчина. Систематическое использование методов теории множеств и теории функций действительного переменного в теории вероятностей, построение основ теории случайных процессов, широкое развитие теории суммирования независимых случайных величин, а также построение нового подхода к задачам статистической физики и стройной системы ее изложения — все это заслуга Александра Яковлевича. Он же разделяет с С. Н. Бернштейном и А. Н. Колмогоровым заслугу создания советской школы теории вероятностей, играющей в современной науке выдающуюся роль. Я счастлив, что мне довелось быть его учеником.

Книжка, написанная нами в период победоносного завершения Великой Отечественной войны, естественно, отражала в рассмотренных нами примерах элементарные постановки военных задач. Теперь, спустя 15 лет после победы, в дни, когда вся страна покрыта лесами новостроек, естественно расширить тематику примеров, иллюстрирующих общие теоретические положения. Именно поэтому, не меняя изложения и элементарного характера книги, я позволил себе заменить на новые большое число примеров. За малыми исключениями те же изменения были внесены мной и во французское издание нашей книжки (Paris, 1960).

Москва,
6 октября 1960 г.

Б. В. Гнеденко

ЧАСТЬ ПЕРВАЯ
ВЕРОЯТНОСТИ

ГЛАВА ПЕРВАЯ
ВЕРОЯТНОСТИ СОБЫТИЙ

§ 1. Понятие вероятности

Когда про какого-нибудь стрелка говорят, что он при данных условиях стрельбы дает 92% попаданий, то это означает, что из сотни выстрелов, произведенных им при некоторых определенных условиях (одна и та же цель на том же расстоянии, та же винтовка и т. д.), *в среднем* бывает примерно 92 удачных (и значит, около 8 неудачных). Конечно, не в каждой сотне будет 92 удачных выстрела; иногда их будет 91 или 90, иногда 93 или 94; подчас число их может оказаться даже заметно меньше или заметно больше чем 92; но *в среднем* при многократном повторении стрельбы в тех же условиях этот процент попаданий будет оставаться неизменным, покуда с течением времени не произойдет каких-нибудь существенных изменений (так, наш стрелок может повысить свое мастерство, доведя средний процент попаданий до 95 и выше). И опыт показывает, что у такого стрелка в большинстве случаев число удачных выстрелов на сотню будет близко к 92; такие сотни, в которых, например, это число меньше чем 88 или больше чем 96, хотя и будут встречаться, но сравнительно редко. Цифра 92%, служащая показателем мастерства нашего стрелка, бывает обычно очень *устойчивой*, т. е. процент попаданий в большинстве стрельб (в тех же условиях) будет для данного стрелка почти один и тот же, лишь в редких, исключительных случаях уклоняясь сколько-нибудь значительно от своего среднего значения.

Рассмотрим еще один пример. На некотором предприятии было замечено, что в данных условиях в среднем 1,6% изготовленных предметов оказываются не удовлетворяющими стандарту и идут в брак. Это означает, что в партии, скажем, из 1000 изделий, еще не подвергнутых браковке, будет примерно 16 непригодных. Иногда, конечно, число бракованных изделий будет несколько больше, иногда несколько меньше, но в среднем это число будет близко к 16, и в большинстве партий по 1000 изделий оно также будет близко к 16. Понятно, что и здесь мы предполагаем условия производства (организация технологического процесса, оборудование, сырье, квалификация рабочих и пр.) неизменными.

Ясно, что таких примеров можно привести сколько угодно. Во всех этих примерах мы видим, что при *однородных массовых* операциях (многократная стрельба, массовое производство изделий и т. п.) процент того или другого вида важных для нас событий (попадание в цель, нестандартность изделия и пр.) при данных условиях почти всегда бывает примерно одним и тем же, лишь в редких случаях уклоняясь сколько-нибудь значительно от некоторой средней цифры. Можно поэтому сказать, что эта средняя цифра является характерным показателем данной массовой операции (при данных строго установленных условиях). Процент попадания описывает нам мастерство стрелка, процент брака оценивает нам доброкачественность продукции. Само собой понятно поэтому, что знание таких показателей очень важно в самых различных областях: военном деле, технике, экономике, физике, химии и пр.; оно позволяет там не только оценивать уже происшедшие массовые явления, но и предвидеть исход той или иной массовой операции в будущем.

Если стрелок в данных условиях стрельбы попадает в цель в среднем 92 из 100 выстрелов, то мы говорим, что для этого стрелка и в этих условиях *вероятность попадания* составляет 92% (или 92/100, или 0,92). Если в данных условиях в среднем на каждую 1000 готовых изделий некоторого предприятия приходится 16 бракованных, то мы говорим, что *вероятность* *изготовления брака* составляет для данного производства 0,016, или 1,6%.

Что же мы вообще называем вероятностью событий в данной массовой операции? На это теперь нетрудно ответить. Массовая операция всегда состоит из повторения большого числа подобных между собою единичных операций (стрельба — из отдельных выстрелов, массовое производство — из изготовления отдельных предметов и т. п.). Нас интересует определенный результат единичной операции (попадание при единичном выстреле, нестандартность отдельного изделия и т. д.) и прежде всего — число таких результатов в той или другой массовой операции (сколько выстрелов попадет в цель, сколько изделий будет забраковано и т. д.). Процент (или вообще долю) таких «удачных» *) результатов в данной массовой операции мы и называем *вероятностью* этого важного для нас результата. При этом всегда надо иметь в виду, что вопрос о вероятности того или другого события (результата) имеет смысл только в точно определенных условиях, в которых протекает наша массовая операция. Всякое существенное изменение этих условий влечет за собой, как правило, изменение интересующей нас вероятности.

Если массовая операция такова, что событие A (например, попадание в цель) наблюдается в среднем a раз среди b единичных операций (выстрелов), то вероятность события A в данных условиях составляет $\frac{a}{b}$ $\left(\text{или } \frac{100a}{b}\,\%\right)$. Можно сказать поэтому, что *вероятностью* «удачного» исхода единичной операции мы называем *отношение* в среднем наблюдающегося *числа таких «удачных» исходов к числу всех единичных операций*, составляющих данную массовую операцию. Само собой разумеется, что если вероятность какого-либо события равна $\frac{a}{b}$, то в каждой серии из b

---

*) Во втором примере скорее следовало бы сказать «неудачных». Однако в теории вероятностей принято называть «удачными» те результаты, которые приводят к осуществлению события, интересующего нас в задаче.

единичных операций это событие может наступить и более чем a раз, и менее чем a раз; только в *среднем* оно наступает примерно a раз; и в большинстве таких серий из b операций число наступлений события A будет близко к a, *в особенности, если b — большое число.*

Пример 1. В некотором городе в течение первого квартала родились:

в январе — 145 мальчиков и 135 девочек,
в феврале — 142 мальчика и 136 девочек,
в марте — 152 мальчика и 140 девочек.

Как велика вероятность рождения мальчика? Доля рождения мальчиков:

$$\text{в январе: } \frac{145}{280} \approx 0{,}518 = 51{,}8\%,$$

$$\text{в феврале: } \frac{142}{278} \approx 0{,}511 = 51{,}1\%,$$

$$\text{в марте: } \frac{152}{292} \approx 0{,}520 = 52{,}0\%.$$

Мы видим, что это среднее арифметическое долей за отдельные месяцы близко к числу $0{,}516 = 51{,}6\%$; искомая вероятность в данных условиях составляет примерно 0,516, или 51,6%. Эта цифра хорошо известна в демографии — науке, изучающей динамику населения; оказывается, что доля рождения мальчика в обычных условиях в различные периоды времени не будет значительно отклоняться от этой цифры.

Пример 2. В начале прошлого века было открыто замечательное явление, получившее название (по имени открывшего его английского ботаника Броуна) броуновского движения. Это явление заключается в том, что мельчайшие частицы вещества, взвешенные в жидкости *), находятся в хаотическом движении, совершающемся без всяких видимых причин.

Долго не могли выяснить причину этого, казалось бы, самопроизвольного движения, пока кинетическая

---

*) То есть находящиеся в состоянии безразличного равновесия.

теория газов не дала простого и исчерпывающего объяснения: движение частиц, взвешенных в жидкости, есть результат ударов молекул жидкости об эти частицы. Кинетическая теория газов дает возможность подсчитать вероятность того, что в данном объеме жидкости не будет ни одной частицы взвешенного вещества, вероятность того, что таких частиц будет одна, две, три и т. д. С целью проверки результатов теории был произведен ряд опытов.

Мы приведем результаты 518 наблюдений шведского физика Сведберга над мельчайшими частицами золота, взвешенными в воде. Было найдено, что в подвергшейся наблюдению части пространства 112 раз не наблюдалось ни одной частицы, 1 частица наблюдалась 168 раз, 2 частицы — 130 раз, 3 частицы — 69 раз, 4 частицы — 32 раза, 5 частиц — 5 раз, 6 частиц — 1 раз, 7 частиц — 1 раз.

Таким образом,

доля 0 частиц равна $\frac{112}{518}=0{,}216$;

» 1 наблюдавшейся частицы равна $\frac{168}{518}=0{,}325$;

» 2 наблюдавшихся частиц » $\frac{130}{518}=0{,}251$;

» 3 » » » $\frac{69}{518}=0{,}133$;

» 4 » » » $\frac{32}{518}=0{,}062$;

» 5 » » » $\frac{5}{518}=0{,}010$;

» 6 » » » $\frac{1}{518}=0{,}002$;

» 7 » » » $\frac{1}{518}=0{,}002$.

Результаты наблюдений, как оказалось, очень хорошо совпали с теоретически предсказанными вероятностями.

Пример 3. В ряде практически важных задач существенно знание того, как часто могут встречаться в тексте те или иные буквы русского алфавита. Так, например, при формировании типографских касс нерационально запасать всех букв по одинаковому числу, так как одни буквы в тексте встречаются значительно чаще, чем другие. Поэтому стремятся, чтобы чаще встречающиеся буквы были представлены в бо́льшем числе. Исследования, произведенные над литературными текстами, привели к оценке частоты букв русского алфавита, включая пробелы между словами, которая сведена в следующую табличку *) (составлена в порядке уменьшения относительной частоты появления):

Буква	Пробел	о	е, ё	а	и	т	н
Относит. частота	0,175	0,090	0,072	0,062	0,062	0,053	0,053
Буква	с	р	в	л	к	м	д
Относит. частота	0,045	0,040	0,038	0,035	0,028	0,026	0,025
Буква	п	у	я	ы	з	ь, ъ	б
Относит. частота	0,023	0,021	0,018	0,016	0,016	0,014	0,014
Буква	г	ч	й	х	ж	ю	ш
Относит. частота	0,013	0,012	0,010	0,009	0,007	0,006	0,006
Буква	ц	щ	э	ф			
Относит. частота	0,004	0,003	0,002	0,002			

Таким образом, исследования показывают, что в среднем из 1000 наудачу выбранных в тексте промежутков и букв на двух местах будет стоять буква «ф», на двадцати восьми — буква «к», на девяносто — буква «о» и на ста семидесяти пяти окажутся промежутки между словами. Эти данные являются достаточно ценными указаниями для формирования наборных касс.

*) Эта табличка заимствована из превосходной популярной книжки А. М. Яглома и И. М. Яглома «Вероятность и информация», Физматгиз, 1960.

В последние годы подобные исследования, уже не ограничивающиеся только статистикой букв в русских текстах, начинают широко использоваться для выяснения особенностей русского языка, а также литературного стиля различных авторов. Данные относительно телеграфных сообщений могут быть использованы для создания наиболее экономичных телеграфных кодов, которые позволяли бы передавать сообщения посредством меньшего числа знаков и тем самым быстрее. Выяснилось, что применяемые сейчас телеграфные коды не являются достаточно экономными.

§ 2. Невозможные и достоверные события

Вероятность события, очевидно, всегда есть положительное число или нуль. Она не может быть больше единицы, потому что у дроби, которой она определяется, числитель не может быть больше знаменателя (число «удачных» операций не может быть больше числа всех предпринятых операций). Условимся обозначать через $\mathsf{P}(A)$ вероятность события A. Каково бы ни было это событие,

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

Чем больше $\mathsf{P}(A)$, тем чаще наступает событие A. Например, чем больше у стрелка вероятность попадания в цель, тем чаще у него удачные выстрелы, тем выше его мастерство. Если вероятность события очень мала, то оно наступает редко; если $\mathsf{P}(A) = 0$, то событие A либо никогда не наступает, либо наступает крайне редко, так что практически можно считать его невозможным. Напротив, если $\mathsf{P}(A)$ близко к единице, то у дроби, которой выражается эта вероятность, числитель близок к знаменателю, т. е. подавляющее большинство операций «удачно»; такое событие наступает в большинстве случаев. Если $\mathsf{P}(A) = 1$, то событие A наступает всегда или почти всегда, так что практически можно считать его, как говорят, «достоверным», т. е. наверняка рассчитывать на его наступление. Если $\mathsf{P}(A) = 1/2$, то событие A наступает примерно в половине всех случаев; это значит, что «удачных» операций наблюдается примерно столько же, сколько «неудачных». Если $\mathsf{P}(A) > 1/2$ событие A наступает чаще, чем не наступает; при $\mathsf{P}(A) < 1/2$ мы имеем обратное явление.

Как мала должна быть вероятность события, чтобы мы практически могли считать его невозможным? На этот вопрос нельзя дать общего ответа, потому что все зависит от того, насколько важно событие, о котором идет речь. Так 0,01 — число небольшое Если мы имеем партию снарядов, и 0,01 есть вероятность того, что снаряд при падении не разорвется, то это означает, что примерно 1% выстрелов останется без результата. С этим можно примириться. Но если 0,01 есть вероятность того, что при прыжке парашют не раскроется, то с этим мириться, конечно, никак нельзя, ибо это означает что в одном из сотни прыжков будет бесцельно гибнуть ценная жизнь бойца-парашютиста. Эти примеры показывают, что в каждой отдельной задаче нужно заранее на основании практических соображений установить, как мала должна быть вероятность события, чтобы мы без ущерба для дела могли не считаться с его возможностью.

§ 3. Задача

Задача. Один стрелок дает 80% попаданий в цель, а другой (при тех же условиях стрельбы) 70%. Найти вероятность поражения цели, если оба стрелка стреляют в нее одновременно. Цель считается пораженной при попадании в нее хотя бы одной из двух пуль.

Первый способ решения. Допустим, что производится 100 двойных выстрелов. Примерно в 80 из них цель будет поражена первым стрелком. Остается около 20 выстрелов, в которых этот стрелок даст промах. Так как второй стрелок поражает цель в среднем 70 раз из 100 выстрелов и, значит, 7 раз из 10 выстрелов, то мы можем ожидать, что в тех 20 выстрелах, в которых первый стрелок даст промах, второму удастся поразить цель примерно 14 раз. Таким образом, при всей сотне выстрелов цель окажется пораженной примерно $80 + 14 = 94$ раза. Вероятность поражения цели при одновременной стрельбе наших двух стрелков равна поэтому 94%, или 0,94.

Второй способ решения. Допустим опять, что производится 100 двойных выстрелов. Мы уже видели, что при этом первый стрелок даст примерно 20 промахов. Так как второй стрелок на сотню выстрелов дает примерно 30 промахов и, значит, на десяток примерно 3 промаха, то можно ожидать, что среди тех 20 выстрелов, в которых промахнется первый стрелок, будет примерно 6 таких, в которых промахнется также и второй. При каждом из этих 6 выстрелов цель останется непораженной, а при каждом из остальных 94 выстрелов по крайней мере один из стрелков выстрелит удачно, и цель будет поражена. Мы снова приходим к выводу, что при двойной стрельбе цель будет поражаться примерно в 94 случаях из 100, т. е. что вероятность поражения составит 94%, или 0,94.

Рассмотренная задача очень проста. Но тем не менее она уже приводит нас к очень важному выводу: могут быть такие случаи, когда полезно уметь находить по вероятностям одних событий вероятности других, более сложных событий. На самом деле таких случаев бывает очень много, и не только в военном деле, а и во всякой науке и всякой практической деятельности, где мы встречаемся с массовыми явлениями.

Конечно, было бы очень неудобно для каждой новой встретившейся задачи такого рода изыскивать свой особый способ решения. Наука всегда старается создать общие правила, знание которых позволило бы уже механически или почти механически решать отдельные похожие друг на друга задачи.

В области массовых явлений наука, которая берет на себя составление таких общих правил, называется *теорией вероятностей*. В этой книжке будут даны первые основы этой науки.

Теория вероятностей есть одна из глав математической науки, подобно арифметике или геометрии. Поэтому путь ее есть путь точного рассуждения, а орудиями ее служат формула, таблица, чертеж и т. п.

ГЛАВА ВТОРАЯ

ПРАВИЛО СЛОЖЕНИЯ ВЕРОЯТНОСТЕЙ

§ 4. Вывод правила сложения вероятностей

Самым простым и самым важным общим правилом, употребляемым при расчете вероятностей, является *правило сложения*, которое мы теперь рассмотрим.

При стрельбе по мишени, изображенной на рис. 1, с данного расстояния для каждого стрелка имеется та или другая вероятность попасть в каждую из областей *1*, *2*, *3*, *4*, *5*, *6*. Пусть для какого-нибудь стрелка вероятность попасть в область *1* составляет 0,24, а вероятность попасть в область *2* — 0,17. Как мы уже знаем, это означает, что из сотни пуль, выпущенных этим стрелком, 24 пули попадают (в среднем) в область *1* и 17 пуль — в область *2*.

Пусть в некотором состязании выстрел признается «отличным», если пуля попала в область *1*, и «хорошим», если она попала в область *2*. Какова вероятность того, что выстрел нашего стрелка будет либо хорошим, либо отличным?

На этот вопрос легко ответить. Из каждой сотни пуль, выпущенных стрелком, примерно 24 ложатся в область *1* и примерно 17 — в область *2*. Значит, в каждой сотне пуль будет примерно $24 + 17 = 41$ пуля, которые попадут либо в область *1*, либо в область *2*. Искомая вероятность равна поэтому $0{,}41 = 0{,}24 + 0{,}17$. Вероятность того, что выстрел будет либо отличным, либо хорошим, равна, следовательно, сумме вероятностей отличного и хорошего выстрелов.

Рассмотрим еще другой пример. Пассажир ждет трамвая № 26 или № 16 возле остановки, у которой останавливаются трамваи четырех маршрутов: №№ 16, 22, 26, 31. Считая, что трамваи всех маршрутов появляются в среднем одинаково часто, найти вероятность того, что первый подошедший к остановке трамвай будет нужного пассажиру маршрута.

Ясно, что вероятность того, что первым подойдет к остановке трамвай № 16, равна $1/4$; такая же вероятность того, что первым подойдет трамвай № 26. Искомая же вероятность, очевидно, равна $1/2$. Но

$$\frac{1}{2}=\frac{1}{4}+\frac{1}{4};$$

поэтому мы можем сказать, что вероятность появления первым трамвая № 16 или № 26 равна сумме вероятностей появления трамвая № 16 и трамвая № 26.

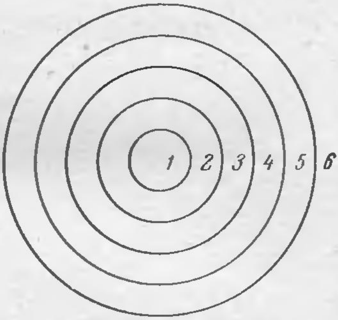

Рис. 1.

Мы можем теперь провести общее рассуждение.

При проведении некоторой массовой операции установлено, что в каждой серии из b единичных операций наблюдается в среднем

a_1 раз некоторый результат A_1,

a_2 » » » A_2,

a_3 » » » A_3

и так далее. Иначе говоря,

вероятность события A_1 равна $\frac{a_1}{b}$,

» » A_2 » $\frac{a_2}{b}$,

» » A_3 » $\frac{a_3}{b}$

и так далее.

Как велика вероятность того, что в некоторой единичной операции наступит какой-либо один (все равно какой) из результатов $A_1, A_2, A_3, \ldots$?

Интересующее нас событие можно назвать (A_1, либо A_2, либо A_3, либо ...)*). Оно в серии из b операций наступает $a_1 + a_2 + a_3 + \ldots$ раз; значит, искомая вероятность равна

$$\frac{a_1 + a_2 + a_3 + \ldots}{b} = \frac{a_1}{b} + \frac{a_2}{b} + \frac{a_3}{b} + \ldots;$$

это можно записать следующей формулой:

$$\mathrm{P}(A_1, \text{ либо } A_2, \text{ либо } A_3, \text{ либо } \ldots) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \mathrm{P}(A_3) + \ldots$$

При этом, как в примерах, так и в общем рассуждении мы все время предполагали, что любые два из рассматриваемых результатов (например, A_1 и A_2) *несовместимы между собою*, т. е. не могут наблюдаться в одной и той же единичной операции. Так, подошедший трамвай не может быть одновременно нужного и ненужного маршрута — он либо удовлетворяет потребность пассажира, либо нет.

Предположение о взаимной несовместимости отдельных интересующих нас результатов очень важно; без него правило сложения становится неверным, и применение его приводит к грубым ошибкам. Рассмотрим, например, задачу, решенную в конце предыдущего параграфа (стр. 14). Там как раз отыскивалась вероятность того, что при двойном выстреле попадет в цель либо первый, либо второй стрелок, причем для первого стрелка вероятность попадания равна 0,8, а для второго 0,7. Если бы мы хотели применить к решению этой задачи правило сложения, то сразу нашли бы, что искомая вероятность равна $0{,}8 + 0{,}7 = 1{,}5$ — результат явно нелепый, так как вероятность события не может быть больше единицы. К такому неверному и бессмысленному ответу мы пришли потому, что применили правило сложения к такому

*) Многоточие здесь и в других подобных случаях означает «и так далее».

случаю, где его применять нельзя: те два результата, о которых идет речь в этой задаче, *совместимы* друг с другом, ибо вполне возможно, что оба стрелка поразят цель при одном и том же двойном выстреле. Значительная часть ошибок, которые делают начинающие при расчете вероятностей, основывается именно на таком неправильном применении правила сложения; необходимо поэтому тщательно остерегаться этой ошибки и при каждом применении правила сложения проверять, действительно ли среди тех событий, к которым мы его хотим применить, каждые два несовместимы друг с другом.

Мы можем теперь дать общую формулировку правила сложения.

Правило сложения. *Вероятность наступления в некоторой операции какого-либо одного* (безразлично какого именно) *из результатов* $A_1, A_2, \ldots, A_n$ *равна сумме вероятностей этих результатов, если каждые два из них несовместимы между собой.*

§ 5. Полная система событий

В Третьем государственном займе восстановления и развития народного хозяйства в течение двадцатилетнего срока его действия третья часть облигаций выигрывает, а остальные две трети выходят в тираж и погашаются по нарицательной стоимости. Иначе говоря, для этого займа каждая облигация имеет вероятность выигрыша, равную 1/3, и вероятность выхода в тираж, равную 2/3. Выигрыш и выход в тираж — *противоположные события*, т. е. такие два события, из которых одно и только одно обязательно наступает для каждой облигации. Сумма их вероятностей

$$\frac{1}{3}+\frac{2}{3}=1,$$

и это не случайно. Вообще, если A_1 и A_2 — два противоположных событий и если в серии из b операций событие A_1 наступает a_1 раз, а событие A_2 — a_2 раз, то, очевидно, $a_1 + a_2 = b$. Но

$$\mathsf{P}(A_1) = \frac{a_1}{b}, \qquad \mathsf{P}(A_2) = \frac{a_2}{b},$$

так что

$$\mathsf{P}(A_1) + \mathsf{P}(A_2) = \frac{a_1}{b} + \frac{a_2}{b} = \frac{a_1 + a_2}{b} = 1.$$

Можно получить тот же результат и из правила сложения: так как противоположные события несовместимы между собою, то

$$\mathsf{P}(A_1) + \mathsf{P}(A_2) = \mathsf{P}(A_1 \text{ либо } A_2);$$

но событие (A_1 либо A_2) есть событие достоверное, так как из определения противоположных событий следует, что оно наверняка должно наступить; поэтому вероятность его равна единице, и мы снова получаем

$$\mathsf{P}(A_1) + \mathsf{P}(A_2) = 1.$$

Сумма вероятностей двух противоположных событий равна единице.

Это правило допускает весьма важное обобщение, которое можно доказать тем же способом. Пусть имеется n (любое число) событий A_1 A_2, ..., A_n, таких, что в каждой единичной операции обязательно должно наступить одно и только одно из этих событий; условимся такую группу событий называть *полной системой*; в частности, всякая пара противоположных событий, очевидно, образует полную систему.

Сумма вероятностей событий, образующих полную систему, равна единице.

Действительно, по определению полной системы любые два из событий этой системы несовместимы между собой, так что правило сложения дает:

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

Но правая часть этого равенства есть вероятность достоверного события и потому равна единице; таким образом, для полной системы

$$\mathsf{P}(A_1)+\mathsf{P}(A_2)+\ \ldots\ +\mathsf{P}(A_n)=1,$$

что и требовалось доказать.

Пример 1. Из каждых 100 выстрелов по мишени, изображенной на рис. 1 (стр. 17), стрелок дает в среднем

44	попадания	в область	*1*,
30	попаданий	» »	*2*,
15	»	» »	*3*,
6	»	» »	*4*,
4	попадания	» »	*5*,
1	попадание	» »	*6*

($44+30+15+6+4+1=100$). Эти шесть результатов стрельбы составляют, очевидно, полную систему событий. Вероятности их соответственно равны

$$0{,}44;\quad 0{,}30;\quad 0{,}15;\quad 0{,}06;\quad 0{,}04;\quad 0{,}01;$$

мы имеем

$$0{,}44+0{,}30+0{,}15+0{,}06+0{,}04+0{,}01=1.$$

Пули, попадающие в область *6*, полностью или частично вовсе не попадают в мишень и не могут быть подсчитаны; это не мешает, однако, найти вероятность попадания в эту область, для чего достаточно вычесть из единицы сумму вероятностей попадания во все другие области.

Пример 2. Статистика показывает, что на некоторой ткацкой фабрике из каждой сотни остановок ткацкого станка, требующих последующей работы ткачихи, в среднем

22	происходят	из-за	обрыва нитей	основы,
31	»	»	» »	утка̀,
27	»	»	смены челнока,	
3	»	»	поломки погонялок,	

а остальные остановки из-за прочих причин.

Мы видим, что кроме прочих причин для остановок станка имеются четыре определенные причины, вероятности которых соответственно равны:

$$0{,}22;\quad 0{,}31;\quad 0{,}27;\quad 0{,}03.$$

Сумма этих вероятностей равна 0,83. Вместе с прочими причинами указанные причины остановок станка составляют полную систему событий; поэтому вероятность остановки станка от прочих причин равна $1 - 0{,}83 = 0{,}17$.

§ 6. Примеры

На установленной теореме о полной системе событий часто с успехом основывают так называемый «априорный», т. е. доопытный расчет вероятностей. Пусть, например, изучается попадание космических частиц на маленькую площадку прямоугольной формы (рис. 2), которая разбита на шесть равных между собой квадратов, занумерованных на рисунке. Интересующие нас площадки находятся в одинаковых условиях, и поэтому нет никаких оснований предполагать, что на какой-нибудь из этих шести квадратов частицы будут попадать чаще, чем на другой. Допустим поэтому, что в каждый из шести квадратов частицы будут в среднем попадать одинаково часто, т. е. что вероятности p_1, p_2, p_3, p_4, p_5, p_6 попадания частиц в эти квадраты равны между собой. Если предположить, что мы интересуемся только частицами, которые попадают лишь на эту площадку, то отсюда будет следовать, что каждое из чисел p равно $1/6$, так как числа эти равны между собой и в сумме дают единицу в силу доказанной выше теоремы. Конечно, сделанный вывод, основанный на ряде допущений, требует для своего подтверждения опытной проверки; однако мы настолько привыкли в подобных случаях к положительному исходу опыта, что с полным практическим основанием полагаемся на наши теоретические допущения и до их опытной проверки. Обычно в таких случаях говорят, что данная операция может иметь n различных *равновероятных* между собой результатов (так, в нашем примере попадание космической частицы на площадку, изображенную на рис. 2, может иметь своим результатом попадание в один из шести квадратов).

Вероятность каждого отдельного из этих n результатов равна $\frac{1}{n}$. Важность такого рода «априорных» расчетов заключается в том, что во многих случаях они позволяют предвидеть вероятность события в таких условиях, где постановка массовых операций либо вовсе невозможна, либо чрезвычайно затруднена.

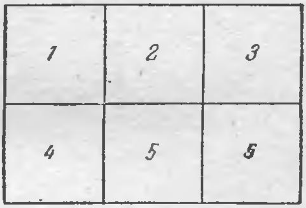

Рис. 2.

Пример 1. В облигациях наших государственных займов номера серий обычно выражаются пятизначными числами. Найти вероятность того, что последняя цифра наудачу взятой выигравшей серии равна 7 (как, например, в серии № 59607). Для этого согласно нашему определению вероятности необходимо рассмотреть длинный ряд тиражных таблиц и подсчитать, сколько выигравших серий имеют номера, оканчивающиеся цифрой 7; отношение этого числа к общему числу выигравших серий и будет искомою вероятностью. Однако каждая из десяти цифр 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 имеет столько же оснований оказаться на последнем месте в номере выигравшей серии, как и любая другая. Поэтому без всяких колебаний высказываем предположение, что искомая вероятность равна 0,1. Правильность этого теоретического «предвидения» легко проверяется: произведя все необходимые подсчеты в пределах какой-нибудь одной тиражной таблицы, можно убедиться, что действительно любая из 10 цифр будет стоять на последнем месте примерно в $\frac{1}{10}$ всех случаев.

Пример 2. Телефонная линия, соединяющая два пункта A и B, отстоящих друг от друга на расстоянии

2 *км*, порвалась в неизвестном месте. Чему равна вероятность того, что она порвалась не далее чем в 450 *м* от пункта A? Разложив мысленно всю линию на отдельные метры, мы можем, в силу реальной однородности всех этих участков, допустить, что вероятность обрыва одна и та же для каждого метра. Отсюда, подобно предыдущему, легко находим, что искомая вероятность равна

$$\frac{450}{2000} = 0{,}225.$$

ГЛАВА ТРЕТЬЯ

УСЛОВНЫЕ ВЕРОЯТНОСТИ И ПРАВИЛО УМНОЖЕНИЯ

§ 7. Понятие условной вероятности

Электрические лампочки производятся на двух заводах, причем первый из них поставляет 70%, а второй 30% всей потребляемой продукции. Из каждых ста лампочек первого завода в среднем 83 стандартных *), а из ста лампочек второго завода — лишь 63 стандартных.

Как легко подсчитать из этих данных, в среднем на каждые 100 электрических лампочек, приобретаемых потребителем, приходится 77 стандартных **), и, следовательно, вероятность купить стандартную лампочку равна 0,77. Но допустим теперь, что мы выяснили, что лампочки, имеющиеся в магазине, изготовлены первым заводом. Тогда вероятность того, что лампочка стандартная, изменится — она будет равна

$$\frac{83}{100} = 0{,}83.$$

Рассмотренный пример показывает, что добавление к общим условиям, в которых происходит операция (в нашем примере — покупка лампочки), некоторого существенно нового условия (у нас в примере —

---

*) Мы станем называть лампочку стандартной (удовлетворяющей стандарту), если она способна прогореть не менее 1200 часов; в противном случае мы назовем лампочку нестандартной.

**) Действительно, $0{,}7 \cdot 83 + 0{,}3 \cdot 63 = 77$.

выяснения того, каким заводом изготовлена лампочка) может изменить вероятность того или иного исхода единичной операции. Это и понятно: ведь само определение понятия вероятности требует, чтобы совокупность условий, при которых проводится данная массовая операция, была точно определена. Присоединяя к этой совокупности какое-либо новое условие, мы, вообще говоря, существенно изменяем эту совокупность. Наша массовая операция проводится после этого уже в новых условиях; фактически, это — уже другая операция, а потому и вероятность того или иного результата в ней уже не та, что в первоначальных условиях.

Мы имеем, таким образом, две различные вероятности одного и того же события — покупки стандартной лампочки, но эти вероятности рассчитаны в различных условиях: покуда мы не налагаем добавочного условия (не учитываем, где изготовлена лампочка), мы имеет *безусловную вероятность* покупки стандартной лампочки, равную 0,77; при наложении же добавочного условия (того, что лампочка изготовлена первым заводом) мы получаем *условную вероятность* 0,83, несколько отличающуюся от предыдущей. Если мы обозначим через A событие — покупку стандартной лампочки и через B событие — выпуск ее первым заводом, то обычно обозначают через $\mathsf{P}(A)$ безусловную вероятность события A и через $\mathsf{P}_B(A)$ вероятность того же события при условии, что состоялось событие B, т. е. что лампочка выпущена первым заводом. Мы имеем, таким образом,

$$\mathsf{P}(A) = 0{,}77; \qquad \mathsf{P}_B(A) = 0{,}83.$$

Так как о вероятности того или иного результата данной операции можно говорить лишь при некоторых точно определенных условиях, то, строго говоря, всякая вероятность есть условная вероятность; безусловных вероятностей в буквальном смысле этого слова существовать не может. Однако в большинстве конкретных задач дело обстоит так, что в основе всех рассматриваемых в данной задаче операций лежит некоторая определенная совокупность условий $\underline{K}$, которые предполагаются выполненными для всех операций. Если при вычислении какой-либо вероятности никаких иных условий, кроме совокупности K, не налагается, то такую вероятность мы назовем *безусловной*; *условной* же будет называться вероятность, вычисленная в предположении, что, кроме общей для всех операций совокупности условий K, выполняются еще те или другие, точно оговоренные дополнительные условия.

Так, в нашем примере мы предполагаем, конечно, что изготовление лампочки производится при некоторых определенных условиях, одних и тех же для всех выпускаемых в продажу лампочек. Это предположение настолько неизбежно и самоочевидно, что мы в формулировке задачи о нем даже не нашли нужным упомянуть. Если на данную лампочку мы не налагаем никаких дополнительных условий, то вероятность того или другого результата испытания лампочки мы называем безусловной. Если же сверх этих условий мы налагаем еще какие-либо дополнительные требования, то вычисляемые при этих требованиях вероятности будут уже условными.

Пример 1. В задаче, описанной нами в начале настоящего параграфа, вероятность выпуска лампочки вторым заводом, очевидно, равна 0,3. Установлено, что лампочка стандартного качества. Какова после этого наблюдения вероятность того, что эта лампочка изготовлена на втором заводе?

Из каждых 1000 выпущенных в продажу лампочек в среднем 770 лампочек имеют стандартное качество, причем из них 581 лампочка первого завода и 189 лампочек второго завода *). После сделанного наблюдения вероятность выпуска лампочки вторым заводом составляет поэтому

$$\frac{189}{770} \approx 0{,}245.$$

*) Это легко подсчитать следующим способом. Из каждых 1000 лампочек 700 в среднем изготовлены первым заводом, а из каждых 100 лампочек первого завода в среднем 83 имеют стандартное качество. Следовательно, из 700 лампочек первого завода стандартное качество будут иметь в среднем $7 \cdot 83 = 581$. Остальные 189 лампочек стандартного качества изготовлены вторым заводом.

Это — условная вероятность выпуска лампочки вторым заводом, вычисленная в предположении, что данная лампочка стандартна. В наших прежних обозначениях мы можем написать

$$\mathrm{P}(\bar{B}) = 0{,}3; \qquad \mathrm{P}_A(\bar{B}) \approx 0{,}245$$

(событие $\bar{B}$ обозначает ненаступление события B).

Пример 2. Производившиеся в некотором районе многолетние наблюдения показали, что из 100 000 детей, достигших десятилетнего возраста, до 40 лет доживает в среднем 82 277, а до 70 лет — 37 977. Найти вероятность того, что если человек достигнет сорокалетнего возраста, то он доживет и до 70 лет.

Так как из 82 277 сорокалетних до 70 лет доживает в среднем 37 977, то вероятность сорокалетнему дожить до семидесяти лет равна

$$\frac{37\,977}{82\,277} \approx 0{,}46.$$

Если через A и B обзначить события, состоящие: первое — в доживании десятилетнего ребенка до 70 лет, а второе — в достижении им 40-летнего возраста, то, очевидно, мы имеем

$$\mathrm{P}(A) = 0{,}37\,977 \approx 0{,}38; \qquad \mathrm{P}_B(A) \approx 0{,}46.$$

§ 8. Вывод правила умножения вероятностей

Вернемся к первому примеру предыдущего параграфа. Из каждой 1000 продаваемых лампочек вторым заводом изготовляется в среднем 300 лампочек, а из этих 300 стандартными оказываются в среднем 189. Отсюда мы получаем, что вероятность изготовления лампочки вторым заводом (событие $\bar{B}$) равна

$$\mathrm{P}(\bar{B}) = \frac{300}{1000} = 0{,}3,$$

а вероятность ее стандартного качества при условии, что она изготовлена вторым заводом, равна

$$\mathrm{P}_{\bar{B}}(A) = \frac{189}{300} = 0{,}63.$$

Так как из каждой 1000 лампочек 189 изготовлены вторым заводом и в то же время имеют стандартное качество, то вероятность совместного наступления событий A и $\bar{B}$ равна

$$\mathsf{P}(A \text{ и } \bar{B}) = \frac{189}{1000} = \frac{300}{1000} \cdot \frac{189}{300} = \mathsf{P}(\bar{B})\,\mathsf{P}_{\bar{B}}(A).$$

Это «правило умножения» легко распространить и на общий случай. Пусть в каждой серии из n операций результат B наступает в среднем m раз, а в каждой серии из m таких операций, где результат B наблюдался, l раз наступает результат A. Тогда в каждой серии из n операций совместное наступление событий B и A будет наблюдаться в среднем l раз. Таким образом,

$$\mathsf{P}(B) = \frac{m}{n}, \qquad \mathsf{P}_B(A) = \frac{l}{m},$$

$$\mathsf{P}(A \text{ и } B) = \frac{l}{n} = \frac{m}{n} \cdot \frac{l}{m} = \mathsf{P}(B)\,\mathsf{P}_B(A). \tag{1}$$

П р а в и л о у м н о ж е н и я. *Вероятность совместного наступления двух событий равна произведению вероятности первого события на условную вероятность второго, вычисленную в предположении, что первое событие состоялось.*

Разумеется, мы можем называть первым любое из двух данных событий, так что на равных основаниях с формулой (1) можем писать и

$$\mathsf{P}(A \text{ и } B) = \mathsf{P}(A)\,\mathsf{P}_A(B), \tag{1'}$$

откуда получаем важное соотношение

$$\mathsf{P}(A)\,\mathsf{P}_A(B) = \mathsf{P}(B)\,\mathsf{P}_B(A). \tag{2}$$

В нашем примере было

$$\mathsf{P}(A \text{ и } \bar{B}) = \frac{189}{1000}, \qquad \mathsf{P}(A) = \frac{77}{100}, \qquad \mathsf{P}_A(\bar{B}) = \frac{189}{770};$$

формула (1′) подтверждается.

П р и м е р. На некотором предприятии 96% изделий признается пригодными (событие A); из каждой сотни годных изделий в среднем 75 оказываются первого сорта (событие B). Найти вероятность того, что изделие, изготовленное на этом предприятии, окажется первого сорта.

Ищется Р (A и B), так как, для того чтобы изделие было первосортным, надо, чтобы оно было пригодным (событие A) и первого сорта (событие B).

В силу условий задачи

$$\mathrm{P}(A)=0{,}96, \qquad \mathrm{P}_A(B)=0{,}75.$$

Поэтому на основании формулы (1′)

$$\mathrm{P}(A \text{ и } B)=0{,}96\cdot 0{,}75=0{,}72.$$

§ 9. Независимые события

При испытании на крепость двух мотков пряжи, изготовленных на разных машинах, оказалось, что для первого мотка образец некоторой длины выдерживает определенную стандартную нагрузку с вероятностью 0,84, а для второго — с вероятностью 0,78 *). Найти вероятность того, что два образца пряжи, взятых с двух разных мотков, оба в состоянии выдержать стандартную нагрузку.

Обозначим через A событие, состоящее в том, что образец, взятый из первого мотка, выдерживает стандартную нагрузку, и через B аналогичное событие для образца из второго мотка. Так как ищется Р (A и B), то мы применяем правило умножения:

$$\mathrm{P}(A \text{ и } B)=\mathrm{P}(A)\,\mathrm{P}_A(B).$$

Здесь, очевидно, $\mathrm{P}(A)=0{,}84$; но что такое $\mathrm{P}_A(B)$? Согласно общему определению условных вероятностей, это есть вероятность того, что образец пряжи из второго мотка выдержит стандартную нагрузку, если такую нагрузку выдержал образец из первого мотка. Но вероятность события B не зависит от того, произошло или нет событие A, хотя бы потому, что эти испытания мы можем производить одновременно, а образцы пряжи выбираются из совершен-

*) Это означает следующее: из 100 образцов, взятых с первого мотка, в среднем 84 образца выдерживают такую нагрузку, а 16 не выдерживают и разрываются.

ко различных мотков, изготовленных на разных машинах. Практически это означает, что процент испытаний, при которых пряжа из второго мотка выдержит стандартную нагрузку, не зависит от того, какой крепости окажется образец из первого мотка, т. е.

$$\mathsf{P}_A(B) = \mathsf{P}(B) = 0{,}78;$$

отсюда следует, что

$$\mathsf{P}(A \text{ и } B) = \mathsf{P}(A)\,\mathsf{P}(B) = 0{,}84 \cdot 0{,}78 = 0{,}6552.$$

Особенность, отличающая этот пример от всех предыдущих, состоит, как мы видим, в том, что здесь вероятность результата B не изменяется от того, что к общим условиям мы прибавляем требование, чтобы состоялось событие A. Иначе говоря, условная вероятность $\mathsf{P}_A(B)$ равна безусловной вероятности $\mathsf{P}(B)$. В этом случае мы будем кратко говорить, что *событие B не зависит от события A.*

Легко убедиться, что если B не зависит от A, то и A не зависит от B; в самом деле, если $\mathsf{P}_A(B) = \mathsf{P}(B)$, то в силу формулы (2) и $\mathsf{P}_B(A) = \mathsf{P}(A)$, а это и значит, что событие A не зависит от события B. Таким образом, независимость двух событий есть свойство *взаимное.* Мы видим, что для взаимно независимых событий правило умножения получает особенно простой вид:

$$\mathsf{P}(A \text{ и } B) = \mathsf{P}(A)\,\mathsf{P}(B). \tag{3}$$

Подобно тому как при всяком применении правила сложения необходимо предварительно установить взаимную несовместимость данных событий, так и при всяком применении правила (3) необходимо убедиться, что события A и B взаимно независимы. Пренебрежение этим указанием приводит к большому числу ошибок. Если события A и B взаимно зависимы, то формула (3) не верна и должна быть заменена более общей формулой (1) или (1').

Правило (3) легко распространяется на случай, когда ищется вероятность наступления не двух, а трех или более взаимно независимых событий. Пусть, например, мы имеем три взаимно независимых события

A, B и C (это означает, что вероятность каждого из них не зависит от того, наступили или не наступили два других события). Так как события A, B и C взаимно независимы, то по правилу (3)

$$\mathsf{P}(A \text{ и } B \text{ и } C) = \mathsf{P}(A \text{ и } B)\,\mathsf{P}(C);$$

вставляя же сюда вместо $\mathsf{P}(A \text{ и } B)$ выражение этой вероятности из формулы (3), находим:

$$\mathsf{P}(A \text{ и } B \text{ и } C) = \mathsf{P}(A)\,\mathsf{P}(B)\,\mathsf{P}(C), \tag{4}$$

и ясно, что такое же правило имеет место в случае, когда рассматриваемая группа содержит любое число событий, лишь бы эти события были взаимно независимы (т. е. вероятность каждого из них не зависела бы от наступления или ненаступления остальных событий).

Вероятность совместного наступления любого числа взаимно независимых событий равна произведению вероятностей этих событий.

Пример 1. Рабочий обслуживает три станка. Вероятность того, что в течение часа станок не потребует внимания рабочего, равна для первого станка 0,9, для второго 0,8 и для третьего 0,85. Найти вероятность того, что в течение некоторого часа ни один станок не потребует к себе внимания рабочего.

Считая, что станки работают независимо друг от друга, находим по формуле (4), что искомая вероятность равна

$$0{,}9 \cdot 0{,}8 \cdot 0{,}85 = 0{,}612.$$

Пример 2. В условиях примера 1 найти вероятность того, что по крайней мере один из трех станков не потребует внимания рабочего в течение часа.

Речь идет о вероятности вида P (A, или B, или C), и поэтому наша мысль прежде всего устремляется, конечно, к правилу сложения. Однако мы немедленно убеждаемся, что это правило в данном случае неприменимо, так как любые два из трех рассматриваемых событий совместимы друг с другом (ничто не мешает двум станкам проработать спокойно в течение одного и того же часа); да и независимо от этого рассуждения мы сразу видим, что сумма трех данных вероятностей значительно превышает единицу и потому вообще никакой вероятностью быть не может.

Для решения поставленной задачи мы заметим, что вероятность того, что станок потребует внимания рабочего, равна 0,1 для первого станка, 0,2 для второго и 0,15 для третьего. Так как эти три события взаимно независимы, то вероятность того, что осуществятся все эти три события, по правилу (4) равна

$$0,1 \cdot 0,2 \cdot 0,15 = 0,0003.$$

Но события «все три станка потребуют к себе внимания» и «по крайней мере один из трех проработает спокойно», очевидно, представляют собой пару противоположных событий. Поэтому сумма их вероятностей равна единице, и, следовательно, искомая вероятность равна $1 - 0,0003 = 0,9997$. Когда вероятность события столь близка к единице, то это событие можно практически считать достоверным. Это означает, что почти всегда в течение часа по меньшей мере один из трех станков будет работать спокойно.

П р и м е р 3. На испытательном стенде испытываются в определенных условиях 250 приборов. Вероятность того, что в течение часа откажет какой-то определенный из этих приборов, равна 0,004, и эта вероятность одна и та же для всех приборов. Найти вероятность того, что в течение часа откажет хотя бы один из испытываемых приборов.

Для отдельного прибора имеется вероятность

$$1 - 0,004 = 0,996$$

того, что этот прибор не откажет. Вероятность того, что не откажет ни один из двухсот пятидесяти испытываемых приборов, равна по правилу умножения для независимых событий произведению 250 множителей, каждый из которых равен 0,996, т. е. равна $(0,996)^{250}$. А вероятность того, что откажет по меньшей мере один из приборов, равна поэтому

$$1 - (0,996)^{250}.$$

Подробный расчет, который мы здесь приводить не будем, показывает, что это число приблизительно равно $^5/_8$. Таким образом, хотя вероятность отказа каждого из приборов в течение часа и не велика, но при испытании большого числа приборов вероятность отказа хотя бы одного из них становится уже весьма значительной.

То рассуждение, которым мы пользовались в двух последних примерах, легко может быть обобщено и приводит к важному общему правилу. В обоих случаях речь шла о вероятности P (A_1, или A_2, или. . ., или A_n) наступления по меньшей мере одного из нескольких взаимно независимых событий $A_1, A_2, \ldots, A_n$. Если мы обозначим через $\bar{A}_k$ событие, состоящее в том, что A_k не наступает, то события A_k и $\bar{A}_k$ взаимно противоположны, так что

$$\mathsf{P}(A_k)+\mathsf{P}(\bar{A}_k)=1.$$

С другой стороны, события $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n$, очевидно, взаимно независимы, так что

$$\mathsf{P}(\bar{A}_1 \text{ и } \bar{A}_2 \text{ и } \ldots \text{ и } \bar{A}_n)=\mathsf{P}(\bar{A}_1)\,\mathsf{P}(\bar{A}_2)\ldots\mathsf{P}(\bar{A}_n)=$$
$$=[1-\mathsf{P}(A_1)][1-\mathsf{P}(A_2)]\ldots[1-\mathsf{P}(A_n)].$$

Наконец, события (A_1, или A_2, или ..., или A_n) и ($\bar{A}_1$, и $\bar{A}_2$, и ..., и $\bar{A}_n$), очевидно, противоположны друг другу (одно из двух: либо наступает по меньшей мере одно из событий A_k, либо наступают все события $\bar{A}_k$). Поэтому

$$\mathsf{P}(A_1 \text{ или } A_2 \text{ или } \ldots \text{ или } A_n)=$$
$$=1-\mathsf{P}(\bar{A}_1 \text{ и } \bar{A}_2 \text{ и } \ldots \text{ и } \bar{A}_n)=$$
$$=1-[1-\mathsf{P}(A_1)]\cdot[1-\mathsf{P}(A_2)]\ldots[1-\mathsf{P}(A_n)]. \quad (5)$$

Эта важная формула, позволяющая вычислить вероятность наступления *по меньшей мере одного* из событий $A_1, A_2, \ldots, A_n$ по данным вероятностям этих событий, верна тогда и только тогда, когда эти события взаимно независимы. В частном случае, когда все события A_k имеют одну и ту же вероятность p (как мы это имели в примере 3),

$$\mathsf{P}(A_1, \text{ или } A_2, \text{ или } \ldots, \text{ или } A_n) = 1 - (1-p)^n. \quad (6)$$

Пример 4. Вытачивается деталь прибора в виде прямоугольного параллелепипеда. Деталь считается удачной, если длина каждого из ее ребер отклоняется от заданных размеров не более чем на 0,01 *мм*. Вероятность отклонений, превышающих 0,01 *мм*, составляет

$$\text{по длине } p_1 = 0{,}08,$$

$$\text{по ширине } p_2 = 0{,}12,$$

$$\text{по высоте } p_3 = 0{,}1;$$

найти вероятность P непригодности детали.

Для того чтобы деталь оказалась неудачной, нужно по крайней мере в одном направлении иметь уклонение от заданного размера, превышающее 0,01 *мм*; так как обычно эти три события могут считаться взаимно независимыми (ибо они в основном вызываются различными причинами), то для решения задачи можно применить формулу (5); это дает

$$\mathsf{P} = 1 - (1-p_1)(1-p_2)(1-p_3) \approx 0{,}27;$$

следовательно, удачными из каждых 100 деталей окажутся в среднем 73.

ГЛАВА ЧЕТВЕРТАЯ

СЛЕДСТВИЯ ПРАВИЛ СЛОЖЕНИЯ И УМНОЖЕНИЯ

§ 10. Вывод некоторых неравенств

Вернемся снова к примеру с лампочками из предыдущей главы (стр. 25). Введем следующие обозначения событий:

A — лампочка стандартного качества,

$\bar{A}$ — лампочка нестандартного качества,

B — лампочка изготовлена первым заводом,

$\bar{B}$ — лампочка изготовлена вторым заводом.

Очевидно, события A и $\bar{A}$ составляют пару противоположных событий; такую же пару составляют события B и $\bar{B}$.

Если лампочка стандартна (A), то либо она изготовлена первым предприятием (A и B), либо вторым (A и $\bar{B}$); так как два последних события, очевидно, несовместимы друг с другом, то по правилу сложения

$$\mathsf{P}(A) = \mathsf{P}(A \text{ и } B) + \mathsf{P}(A \text{ и } \bar{B}). \tag{1}$$

Тем же путем мы находим:

$$\mathsf{P}(B) = \mathsf{P}(A \text{ и } B) + \mathsf{P}(\bar{A} \text{ и } B). \tag{2}$$

Наконец, рассмотрим событие (A или B); для его наступления имеются, очевидно, следующие три возможности:

$$1)\ A \text{ и } B,$$

$$2)\ A \text{ и } \bar{B},$$

$$3)\ \bar{A} \text{ и } B;$$

из этих трех возможностей любые две несовместимы друг с другом; поэтому по теореме сложения

$$\mathsf{P}(A \text{ или } B) = \mathsf{P}(A \text{ и } B) + \mathsf{P}(A \text{ и } \bar{B}) + \mathsf{P}(\bar{A} \text{ и } B). \quad (3)$$

Складывая почленно равенства (1) и (2) и учитывая равенство (3), мы легко находим:

$$\mathsf{P}(A) + \mathsf{P}(B) = \mathsf{P}(A \text{ и } B) + \mathsf{P}(A \text{ или } B),$$

откуда

$$\mathsf{P}(A \text{ или } B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \text{ и } B). \quad (4)$$

Мы пришли к очень важному результату; при этом, хотя мы и вели рассуждение для частного примера, оно было столь общим, что вывод может считаться установленным для любой пары событий A и B. До сих пор мы получали выражения для вероятности P (A или B) только при весьма частных предположениях относительно связи между событиями A и B (мы предполагали их сначала несовместимыми, а потом — независимыми между собою). Формула (4), которую мы получили теперь, имеет место без всяких дополнительных предположений для любой пары событий A и B. Правда, мы не должны забывать одного существенного различия между формулой (4) и нашими прежними формулами. В прежних формулах вероятность P (A или B) всегда выражалась через одни только вероятности P (A) и P (B), так что, зная вероятности событий A и B, мы всегда могли однозначно определить вероятность события (A или B). В формуле (4) дело обстоит иначе: для вычисления по ней величины P (A или B) необходимо кроме P (A) и P (B) знать еще P (A и B), т. е. вероятность совместного наступления событий A и B; найти же эту вероятность в общем случае, т. е. при произвольной связи между событиями A и B, бывает обычно не легче, чем найти P (A или B); поэтому для непосредственных практических расчетов формулой (4) пользоваться приходится редко, но теоретическое значение ее очень велико.

Убедимся прежде всего, что из формулы (4) в качестве частных случаев легко могут быть получены наши прежние формулы. Если события A и B несовместимы друг с другом, то событие (A и B) невозможно, P (A и B) $= 0$, и формула (4) приводит к соотношению

$$\mathsf{P}(A \text{ или } B) = \mathsf{P}(A) + \mathsf{P}(B),$$

т. е. к правилу сложения. Если события A и B взаимно независимы, то по формуле (3) на стр. 31

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

и формула (4) дает

$$\mathsf{P}(A \text{ или } B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A)\,\mathsf{P}(B) = \\ = 1 - [1 - \mathsf{P}(A)]\,[1 - \mathsf{P}(B)],$$

т. е. формулу (5) на стр. 34 (для случая $n = 2$).

Далее, выведем из формулы (4) одно важное следствие. Так как во всех случаях P (A и B) $\geqslant 0$, то из формулы (4) во всех случаях следует

$$\mathsf{P}(A \text{ или } B) \leqslant \mathsf{P}(A) + \mathsf{P}(B). \qquad (5)$$

Это неравенство легко может быть распространено на любое число событий. Так, например, в случае трех событий в силу (5)

$$\mathsf{P}(A, \text{ или } B, \text{ или } C) \leqslant \mathsf{P}(A \text{ или } B) + \mathsf{P}(C) \leqslant \\ \leqslant \mathsf{P}(A) + \mathsf{P}(B) + \mathsf{P}(C),$$

и тем же путем, очевидно, можно от трех событий перейти к четырем и т. д. Получаем общий результат:

Вероятность наступления по меньшей мере одного из нескольких событий никогда не превышает суммы вероятностей этих событий.

При этом знак равенства имеет место только в том случае, когда любые два из данных событий несовместимы между собой.

§ 11. Формула полной вероятности

Вернемся еще раз к примеру с лампочками на стр. 25 и будем пользоваться для различных результатов испытания обозначениями, введенными на стр. 36. Вероятность стандартного качества лампочки при условии, что она изготовлена вторым заводом, равна, как мы видели уже не раз,

$$\mathrm{P}_{\overline{B}}(A) = \frac{189}{300} = 0{,}63;$$

вероятность того же события при условии, что лампочка изготовлена первым заводом, составляет

$$\mathrm{P}_B(A) = \frac{581}{700} = 0{,}83.$$

Допустим, что эти два числа нам известны и что известны также вероятности изготовления лампочки первым заводом

$$\mathrm{P}(B) = 0{,}7$$

и вторым

$$\mathrm{P}(\overline{B}) = 0{,}3.$$

Требуется найти безусловную вероятность $\mathrm{P}(A)$, т. е. вероятность стандартности отдельной лампочки, вычисленную без всяких предположений о месте ее изготовления.

Чтобы решить эту задачу, будем рассуждать так. Обозначим через E двойное событие, состоящее в том, что 1) лампочка выпущена первым заводом и 2) она стандартна, а через F — такое же событие для второго завода. Так как всякая стандартная лампочка изготовлена первым или вторым заводом, то событие A равносильно событию «E или F», а так как события E и F друг с другом несовместимы, то по правилу сложения

$$\mathrm{P}(A) = \mathrm{P}(E) + \mathrm{P}(F). \tag{6}$$

С другой стороны, чтобы имело место событие E, надо, чтобы 1) лампочка была изготовлена первым заводом (B) и 2) чтобы она была стандартна (A); поэтому событие E равносильно событию (B и A), откуда по правилу умножения

$$\mathrm{P}(E) = \mathrm{P}(B)\,\mathrm{P}_B(A).$$

Совершенно тем же путем находим:

$$\mathsf{P}(F) = \mathsf{P}(\bar{B})\,\mathsf{P}_{\bar{B}}(A),$$

а подставляя эти выражения в равенство (6):

$$\mathsf{P}(A) = \mathsf{P}(B)\,\mathsf{P}_{B}(A) + \mathsf{P}(\bar{B})\,\mathsf{P}_{\bar{B}}(A).$$

Эта формула и решает поставленную нами задачу. Подставляя данные числа, находим $\mathsf{P}(A) = 0{,}77$.

Пример. Для посева заготовлены семена пшеницы сорта I, содержащие небольшое количество примесей других сортов — II, III, IV. Возьмем одно из этих зерен. Событие, состоящее в том, что это зерно сорта I, обозначим через A_1, что оно сорта II — через A_2, сорта III — через A_3 и, наконец, сорта IV — через A_4. Известно, что вероятности того, что наудачу взятое зерно окажется того или иного сорта, равны

$$\mathsf{P}(A_1) = 0{,}96; \quad \mathsf{P}(A_2) = 0{,}01; \quad \mathsf{P}(A_3) = 0{,}02; \quad \mathsf{P}(A_4) = 0{,}01.$$

(Сумма этих четырех чисел равна единице, как это и должно быть для полной системы событий.)

Вероятность того, что из зерна вырастает колос, содержащий не менее 50 зерен, равна:

1) 0,50 из зерна сорта I,
2) 0,15 » » » II,
3) 0,20 » » » III,
4) 0,05 » » » IV.

Требуется найти безусловную вероятность того, что колос будет иметь не менее 50 зерен.

Пусть K — событие, состоящее в том, что колос будет содержать не менее 50 зерен; тогда по условию задачи

$$\mathsf{P}_{A_1}(K) = 0{,}50; \quad \mathsf{P}_{A_2}(K) = 0{,}15;$$
$$\mathsf{P}_{A_3}(K) = 0{,}20; \quad \mathsf{P}_{A_4}(K) = 0{,}05.$$

Нашей задачей является определение $\mathsf{P}(K)$. Обозначим через E_1 событие, состоящее в том, что зерно окажется сорта I и колос, выращенный из него, будет содержать не менее 50 зерен, так что E_1 равносильно событию (A_1 и K); подобным же образом обозначим

через E_2 — событие (A_2 и K),
» E_3 — » (A_3 и K),
» E_4 — » (A_4 и K).

Очевидно, что для наступления события K необходимо, чтобы наступило одно из событий E_1, E_2, E_3, E_4, а так как любые два из этих событий друг с другом несовместимы, то по правилу сложения получаем:

$$\mathsf{P}(K)=\mathsf{P}(E_1)+\mathsf{P}(E_2)+\mathsf{P}(E_3)+\mathsf{P}(E_4). \qquad (7)$$

С другой стороны, по правилу умножения,

$$\mathsf{P}(E_1)=\mathsf{P}(A_1 \text{ и } K)=\mathsf{P}(A_1)\,\mathsf{P}_{A_1}(K),$$
$$\mathsf{P}(E_2)=\mathsf{P}(A_2 \text{ и } K)=\mathsf{P}(A_2)\,\mathsf{P}_{A_2}(K),$$
$$\mathsf{P}(E_3)=\mathsf{P}(A_3 \text{ и } K)=\mathsf{P}(A_3)\,\mathsf{P}_{A_3}(K),$$
$$\mathsf{P}(E_4)=\mathsf{P}(A_4 \text{ и } K)=\mathsf{P}(A_4)\,\mathsf{P}_{A_4}(K).$$

Вставляя эти выражения в равенство (7), мы находим:

$$\mathsf{P}(K)=\mathsf{P}(A_1)\,\mathsf{P}_{A_1}(K)+\mathsf{P}(A_2)\,\mathsf{P}_{A_2}(K)+ \\ +\mathsf{P}(A_3)\,\mathsf{P}_{A_3}(K)+\mathsf{P}(A_4)\,\mathsf{P}_{A_4}(K),$$

что и решает, очевидно, нашу задачу.

Подставляя данные числа, находим:

$$\mathsf{P}(K)=0{,}486.$$

Два примера, которые мы здесь подробно рассмотрели, приводят нас к важному общему правилу; мы можем теперь сформулировать и доказать его без всяких затруднений. Пусть данная операция допускает результаты A_1, A_2, ..., A_n, образующие полную систему событий (напомним: это означает, что любые два из этих событий друг с другом несовместимы и что какое-нибудь из них обязательно должно наступить). Тогда для любого возможного результата K этой операции имеет место соотношение

$$\mathsf{P}(K)=\mathsf{P}(A_1)\,\mathsf{P}_{A_1}(K)+ \\ +\mathsf{P}(A_2)\,\mathsf{P}_{A_2}(K)+\ \ldots\ +\mathsf{P}(A_n)\,\mathsf{P}_{A_n}(K). \qquad (8)$$

Это правило (8) обычно называют *формулой полной вероятности*. Доказательство его проводится в точности так же, как в двух рассмотренных нами примерах; во-первых, наступление событий K требует наступления одного из событий (A_i и K), так что по правилу сложения

$$\mathrm{P}(K) = \sum_{i=1}^{n} \mathrm{P}(A_i \text{ и } K); \tag{9}$$

во-вторых, по правилу умножения,

$$\mathrm{P}(A_i \text{ и } K) = \mathrm{P}(A_i)\,\mathrm{P}_{A_i}(K);$$

вставляя эти выражения в равенство (9), мы и приходим к формуле (8).

§ 12. Формула Байеса

Формулы предыдущего параграфа позволяют нам вывести один важный результат, имеющий многочисленные приложения. Мы начнем с формального вывода, отложив выяснение реального смысла окончательной формулы до рассмотрения примеров.

Пусть снова события $A_1, A_2, \ldots, A_n$ представляют собой полную систему результатов некоторой операции. Если тогда K означает произвольный результат этой операции, то по правилу умножения

$$\mathrm{P}(A_i \text{ и } K) = \mathrm{P}(A_i)\,\mathrm{P}_{A_i}(K) = \mathrm{P}(K)\,\mathrm{P}_K(A_i) \qquad (1 \leqslant i \leqslant n),$$

откуда

$$\mathrm{P}_K(A_i) = \frac{\mathrm{P}(A_i)\,K_{A_i}(K)}{\mathrm{P}(K)} \qquad (1 \leqslant i \leqslant n),$$

или, выражая знаменатель полученной дроби по формуле полной вероятности (8) предыдущего параграфа,

$$\mathrm{P}_K(A_i) = \frac{\mathrm{P}(A_i)\,\mathrm{P}_{A_i}(K)}{\sum\limits_{r=1}^{n} \mathrm{P}(A_r)\,\mathrm{P}_{A_r}(K)} \qquad (1 \leqslant i \leqslant n). \tag{10}$$

Это — *формула Байеса*, имеющая много приложений в практике вычисления вероятностей. Чаще всего приходится применять ее в положениях, которые иллюстрируются следующим примером

Пусть стрельба ведется по цели, расположенной на прямолинейном участке MN (рис. 3), который мы мысленно разбили на пять небольших участков, a, b', b'', c', c''. Допустим, что точное местоположение цели нам неизвестно мы знаем только вероятности того, что

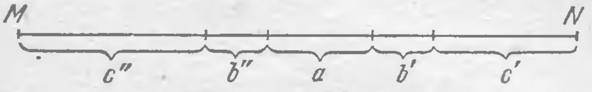

Рис. 3.

цель лежит на том или другом из наших пяти участков; пусть эти вероятности равны

$$\mathsf{P}(a) = 0{,}48;\ \mathsf{P}(b') = \mathsf{P}(b'') = 0{,}21;\ \mathsf{P}(c') = \mathsf{P}(c'') = 0{,}05,$$

где через a, b', b'', c', c'' обозначены следующие события: цель находится в отрезке a, b', b'', c', c'' (сумма этих чисел равна единице). Наибольшая вероятность соответствует отрезку a, куда мы поэтому, естественно, и направляем наш выстрел. Однако из-за неизбежных ошибок стрельбы цель может оказаться пораженной и тогда, когда она находится не в a, а в каком-либо из других отрезков. Пусть вероятность поражения цели (события K) составляет:

$\mathsf{P}_a(K) = 0{,}56$, если цель лежит в отрезке a,
$\mathsf{P}_{b'}(K) = 0{,}18$, » » » » » b',
$\mathsf{P}_{b''}(K) = 0{,}16$, » » » » » b'',
$\mathsf{P}_{c'}(K) = 0{,}06$, » » » » » c',
$\mathsf{P}_{c''}(K) = 0{,}02$, » » » » » c''.

Допустим, что выстрел произведен и цель оказалась пораженной (состоялось событие K). В результате этого вероятности различных положений цели, которые мы имели раньше (т. е. числа $\mathsf{P}(a)$, $\mathsf{P}(b')$, ...), подвергаются переоценке; качественная сторона этой переоценки ясна без всяких вычислений; мы стреляли по отрезку a и попали в цель — ясно, что вероятность $\mathrm{P}(a)$ при этом должна увеличиться; но мы хотим точно количественно учесть произведенную нашим выстрелом переоценку, т. е. мы хотим найти точное выражение вероятностей $\mathrm{P}_K(a)$, $\mathrm{P}_K(b')$, ... различных возможных положений цели при условии, что произведенным выстрелом цель была поражена. Формула Байеса (10) сразу дает нам ответ на этот вопрос:

$$\mathrm{P}_K(a) = \mathrm{P}(a)\,\mathrm{P}_a(K)\,[\mathrm{P}(a)\,\mathrm{P}_a(K) + \mathrm{P}(b')\,\mathrm{P}_{b'}(K) + \\ + \mathrm{P}(b'')\,\mathrm{P}_{b''}(K) + \mathrm{P}(c')\,\mathrm{P}_{c'}(K) + \mathrm{P}(c'')\,\mathrm{P}_{c''}(K)]^{-1} \approx 0{,}8;$$

мы видим, что $\mathrm{P}_K(a)$ действительно больше, чем $\mathrm{P}(a)$.

Подобным же образом легко находим и вероятности $\mathrm{P}_K(b')$, ... других положений цели. Для вычисления полезно заметить, что выражения, даваемые для этих вероятностей формулой Байеса, отличаются друг от друга только своими числителями; знаменатель же у них один и тот же; он равен $\mathrm{P}(K) \approx 0{,}34$.

Общая схема подобного рода положений может быть описана так. Условия операции содержат некоторый неизвестный элемент, относительно которого может быть сделано n различных гипотез: A_1, A_2, ..., A_n, образующих полную систему событий; по тем или другим причинам нам известны вероятности $\mathrm{P}(A_i)$ этих гипотез до испытания; известно также, что гипотеза A_i «сообщает» некоторому событию K (например, попаданию в цель) вероятность $\mathrm{P}_{A_i}(K)$ $(1 \leqslant i \leqslant n)$ ($\mathrm{P}_{A_i}(K)$ есть вероятность события K, вычисленная при условии, что справедлива гипотеза A_i). Если в результате опыта событие K наступило, то это вызывает переоценку вероятностей гипотез A_i и задача состоит в том, чтобы найти новые вероятности $\mathrm{P}_K(A_i)$ этих гипотез; ответ дается формулой Байеса.

В артиллерийской практике производится так называемая пристрелка, имеющая целью уточнить наши знания об условиях стрельбы. При этом неизвестным элементом, требующим уточнения, может служить не только положение цели, но и любой другой элемент условий стрельбы, влияющий на ее эффективность (в частности, та или другая особенность употребляемого оружия).

Очень часто бывает, что таких пробных выстрелов производится не один, а несколько, и ставится задача о вычислении новых вероятностей гипотез на основе полученных результатов стрельбы. Во всех таких случаях формула Байеса также легко решает поставленную задачу.

Для сокращения записи положим в рассмотренной нами общей схеме

$$\mathsf{P}(A_i) = \mathsf{P}_i, \quad \mathsf{P}_{A_i}(K) = p_i \qquad (1 \leqslant i \leqslant n),$$

так что формула Байеса получает простой вид

$$\mathsf{P}_K(A_i) = \frac{\mathsf{P}_i p_i}{\sum\limits_{r=1}^{n} \mathsf{P}_r p_r}.$$

Допустим, что произведено s пробных выстрелов, причем результат K наступил m раз и не наступил $s - m$ раз. Обозначим через K^* полученный результат серии из s выстрелов. Мы можем допустить, что результаты отдельных выстрелов представляют собой взаимно независимые события. Если справедлива гипотеза A_i, вероятность результата K равна p_i и, значит, вероятность противоположного события (т. е. того, что K не наступит) равна $1 - p_i$.

Вероятность того, что результат K наступил при определенных m выстрелах, по правилу умножения для независимых событий равна $p_i^m(1 - p_i)^{s-m}$. Так как эти m выстрелов, при которых наступил результат K, могут оказаться любыми из s произведенных, то событие K^* может осуществиться C_s^m несовместимыми способами. Таким образом, по правилу сложения вероятностей

$$\mathsf{P}_{A_i}(K^*) = C_s^m p_i^m (1 - p_i)^{s-m} \qquad (1 \leqslant i \leqslant n),$$

и формула Байеса дает:

$$\mathsf{P}_{K^*}(A_i)=\frac{P_i p_i^m (1-p_i)^{s-m}}{\sum\limits_{r=1}^{n} P_r p_r^m (1-p_r)^{s-m}} \quad (1\leqslant i\leqslant n), \qquad (11)$$

что и решает поставленную задачу. Понятно, что подобные задачи возникают не только в практике артиллериста, но и в других областях человеческой деятельности.

Пример 1. В задаче, рассмотренной нами в начале настоящего параграфа, найти вероятность того, что цель лежит в отрезке a, если два последовательных выстрела по этому отрезку дали попадание.

Обозначая через K^* двукратное попадание в цель, согласно формуле (11) имеем:

$$\mathsf{P}_{K^*}(a)=\frac{\mathsf{P}(a)\,[\mathsf{P}_a(K)]^2}{\mathsf{P}(a)\,[\mathsf{P}_a(K)]^2+\mathsf{P}(b')\,[\mathsf{P}_{b'}(K)]^2+\dots};$$

предоставляем читателю произвести несложный расчет и убедиться, что в результате двукратного попадания вероятность того, что цель расположена в отрезке a, еще более увеличилась.

Пример 2. Вероятность того, что в некотором производстве изделие удовлетворяет стандарту, равна 0,96. Предлагается упрощенная система испытаний *), которая для изделий, удовлетворяющих стандарту, дает положительный результат с вероятностью 0,98, а для изделий, ему не удовлетворяющих, — лишь с вероятностью 0,05. Какова вероятность, что изделие,

*) Необходимость в упрощенном контроле встречается на практике весьма часто. Например, если бы при выпуске в свет электрических лампочек все они подвергались проверке на способность их горения в течение не менее чем, скажем, 1200 часов, то потребитель получал бы лишь одни перегоревшие или почти перегоревшие лампочки. Приходится испытание на срок горения заменить другим испытанием, например проверкой лампочки на зажигаемость.

дважды выдержавшее упрощенное испытание, удовлетворяет стандарту?

Здесь полная система гипотез состоит из двух противоположных событий: 1) изделие удовлетворяет стандарту, 2) изделие не удовлетворяет стандарту. Вероятности этих гипотез до опыта соответственно равны $P_1 = 0{,}96$ и $P_2 = 0{,}04$. При первой гипотезе вероятность того, что изделие выдержит испытание, равна $p_1 = 0{,}98$, а при второй $p_2 = 0{,}05$. После двукратного опыта вероятность первой гипотезы на основании формулы (11) равна

$$\frac{P_1p_1^2}{P_1p_1^2 + P_2p_2^2} = \frac{0{,}96 \cdot (0{,}98)^2}{0{,}96 \cdot (0{,}98)^2 + 0{,}04 \cdot (0{,}05)^2} \approx 0{,}9999.$$

Мы видим, что если изделие выдержало указанное в условии задачи испытание, то только в одном случае из десяти тысяч мы можем совершить ошибку, считая его стандартным. Это, конечно, вполне удовлетворяет требованиям практики.

П р и м е р 3. При исследовании больного имеется подозрение на одно из трех заболеваний: A_1, A_2, A_3. Их вероятности в данных условиях равны соответственно

$$P_1 = \frac{1}{2}; \quad P_2 = \frac{1}{6}; \quad P_3 = \frac{1}{3}.$$

Для уточнения диагноза назначен некоторый анализ, дающий положительный результат с вероятностью 0,1 в случае заболевания A_1, с вероятностью 0,2 в случае заболевания A_2 и с вероятностью 0,9 в случае заболевания A_3. Анализ был произведен пять раз и дал четыре раза положительный результат и один раз отрицательный. Требуется найти вероятность каждого заболевания после анализа.

В случае заболевания A_1 вероятность указанных исходов анализов равна, по правилу умножения, $p_1 = C_5^4 \cdot (0{,}1)^4 \cdot 0{,}9$. Для второй гипотезы эта вероятность равна $p_2 = C_5^4 \cdot (0{,}2)^4 \cdot 0{,}8$ и для третьей — $p_3 = C_5^4 \cdot (0{,}9)^4 \cdot 0{,}1$.

По формуле Байеса находим, что после анализов вероятность заболевания A_1 оказывается равной

$$\frac{P_1p_1}{P_1p_1+P_2p_2+P_3p_3}=$$

$$=\frac{\frac{1}{2}\cdot(0{,}1)^4\cdot 0{,}9}{\frac{1}{2}\cdot(0{,}1)^4\cdot 0{,}9+\frac{1}{6}\cdot(0{,}2)^4\cdot 0{,}8+\frac{1}{3}\cdot(0{,}9)^4\cdot 0{,}1}\approx 0{,}002;$$

вероятность заболевания A_2 —

$$\frac{P_2p_2}{P_1p_1+P_2p_2+P_3p_3}=$$

$$=\frac{\frac{1}{6}\cdot(0{,}2)^4\cdot 0{,}8}{\frac{1}{2}\cdot(0{,}1)^4\cdot 0{,}9+\frac{1}{6}\cdot(0{,}2)^4\cdot 0{,}8+\frac{1}{3}\cdot(0{,}9)^4\cdot 0{,}1}\approx 0{,}01$$

и вероятность заболевания A_3 —

$$\frac{P_3p_3}{P_1p_1+P_2p_2+P_3p_3}=$$

$$=\frac{\frac{1}{3}\cdot(0{,}9)^4\cdot 0{,}1}{\frac{1}{2}\cdot(0{,}1)^4\cdot 0{,}9+\frac{1}{6}\cdot(0{,}2)^4\cdot 0{,}8+\frac{1}{4}\cdot(0{,}9)^4\cdot 0{,}1}\approx 0{,}988.$$

Так как эти три события A_1, A_2, A_3 и после опыта образуют, очевидно, полную систему событий, то для контроля произведенного расчета можно сложить три полученных числа и убедиться, что сумма их по-прежнему равна единице.

ГЛАВА ПЯТАЯ

СХЕМА БЕРНУЛЛИ

§ 13. Примеры

Пример 1. Среди волокон хлопка определенного сорта в среднем 75% имеют длину, меньшую чем 45 *мм*, и 25% — длину, большую (или равную) 45 *мм*. Найти вероятность того, что среди трех наудачу взятых волокон два будут короче, а одно длиннее 45 *мм*.

Обозначим событие — выбор волокна с длиной, меньшей 45 *мм*, — через A, а событие — выбор волокна с длиной, большей 45 *мм*, — через B; тогда ясно, что

$$\mathsf{P}(A)=\frac{3}{4}; \qquad \mathsf{P}(B)=\frac{1}{4}.$$

Условимся далее схемой AAB обозначать сложное событие: два первых выбранных волокна оказались короче 45 *мм*, а третье длиннее 45 *мм* Ясно, какое значение будут иметь схемы BBA, ABA и т. п. Наша задача состоит в вычислении вероятности события C, состоящего в том, что из трех волокон два окажутся короче 45 *мм*, а одно длиннее 45 *мм*. Очевидно, для этого должна осуществиться одна из следующих схем:

$$AAB, \quad ABA, \quad BAA. \tag{1}$$

Так как любые два из этих трех результатов несовместимы между собой, то по правилу сложения

$$\mathsf{P}(C)=\mathsf{P}(AAB)+\mathsf{P}(ABA)+\mathsf{P}(BAA).$$

Все три слагаемых правой части равны между собой, так как результаты выбора волокон мы можем считать взаимно независимыми событиями. Вероятность каждой из схем (1) по правилу умножения вероятностей для независимых событий представится как произведение трех множителей, из которых два равны $\mathsf{P}(A) = 3/4$, а один — $\mathsf{P}(B) = 1/4$. Таким образом, вероятность каждой из трех схем (1) равна

$$\left(\frac{3}{4}\right)^2 \cdot \frac{1}{4} = \frac{9}{64},$$

и следовательно,

$$\mathsf{P}(C) = 3 \cdot \frac{9}{64} = \frac{27}{64},$$

что и решает поставленную задачу.

Пример 2. В результате наблюдений, продолжавшихся многие десятки лет, найдено, что из каждой тысячи новорожденных в среднем рождается 515 мальчиков и 485 девочек. В некоторой семье шестеро детей. Найти вероятность того, что среди них не больше двух девочек.

Для наступления события, вероятность которого мы ищем, надо, чтобы в семье было либо 0, либо 1, либо 2 девочки. Вероятности этих частных событий мы обозначим соответственно через P_0, P_1, P_2. Очевидно, что по правилу сложения вероятностей искомая вероятность

$$P = P_0 + P_1 + P_2. \tag{2}$$

Для каждого ребенка вероятность того, что это мальчик, равна 0,515 и, значит, вероятность того, что это девочка, равна 0,485.

Проще всего найти P_0; это — вероятность того, что все дети в семье оказались мальчиками. Так как событие — рождение ребенка того или другого пола — мы можем рассматривать как независимое от рождения остальных детей, то по правилу умножения вероятность того, что все шесть детей окажутся мальчиками, равна произведению шести множителей, равных 0,515, т. е.

$$P_0 = (0{,}515)^6 \approx 0{,}018.$$

Перейдем теперь к вычислению P_1, т. е. вероятности того, что среди шести детей в семье один ребенок окажется девочкой, а остальные пять — мальчиками. Это событие может произойти шестью различными способами, смотря по тому, которым ребенком по порядку рождений окажется девочка (первым, вторым и т. д.). Рассмотрим какую-нибудь из разновидностей нашего события, например ту, что девочка родилась четвертым ребенком. Вероятность этой разновидности по правилу умножения равна произведению шести множителей, из которых пять равны 0,515, а шестой (стоящий на четвертом месте) равен 0,485, т. е. эта вероятность равна $(0{,}515)^5 \cdot 0{,}485$. Такова же вероятность и всякой другой из пяти возможных разновидностей интересующего нас события; поэтому вероятность P_1 этого события по правилу сложения равна сумме шести чисел, равных $(0{,}515)^5 \cdot 0{,}485$, т. е.

$$P_1 = 6 \cdot (0{,}515)^5 \cdot 0{,}485 \approx 0{,}105.$$

Обращаемся теперь к вычислению P_2 (вероятность того, что два ребенка окажутся девочками, а четыре — мальчиками). Подобно предыдущему, мы сразу замечаем, что событие это допускает целый ряд разновидностей (одной из разновидностей будет, например, такая: второй и пятый ребенок по порядку рождений оказались девочками, а остальные — мальчиками). Вероятность каждой разновидности по правилу умножения равна $(0{,}515)^4 \cdot (0{,}485)^2$, а следовательно, P_2 по правилу сложения равно числу $(0{,}515)^4 \cdot (0{,}485)^2$, умноженному на число всех разновидностей рассматриваемого типа; к определению этого последнего числа и сводится, таким образом, вся задача.

Каждая разновидность характеризуется тем, что из шести детей двое оказываются девочками, а остальные — мальчиками; число различных разновидностей равно, следовательно, числу различных способов выбора двух детей из шести имеющихся. Число таких выборов равно числу сочетаний из шести предметов по два, т. е.

$$C_6^2 = \frac{6 \cdot 5}{2 \cdot 1} = 15.$$

Таким образом,

$$P_2 = C_6^2 \cdot (0{,}515)^4 \cdot (0{,}485)^2 = 15 \cdot (0{,}515)^4 \cdot (0{,}485)^2 \approx 0{,}247.$$

Сопоставляя полученные результаты, находим:

$$P = P_0 + P_1 + P_2 \approx 0{,}018 + 0{,}105 + 0{,}247 = 0{,}370.$$

Таким образом, немного реже, чем в четырех случаях из десяти (с вероятностью $P \approx 0{,}37$), в таких многодетных семьях будет не более трети девочек и, значит, не менее двух третей мальчиков.

§ 14. Формулы Бернулли

В предыдущем параграфе мы познакомились на ряде примеров со схемой *повторения испытаний*, в каждом из которых может осуществляться некоторое событие A. В слово «испытание» вкладывается весьма богатый и разнообразный смысл. Так, если производят стрельбу по некоторой цели, то под испытанием разумеют каждый отдельный выстрел. Если испытывают электрические лампочки на длительность горения, то под испытанием понимают испытание каждой лампочки. Если изучают состав новорожденных по полу, весу или росту, то под испытанием понимают обследование отдельного младенца. Вообще под испытанием мы будем понимать в дальнейшем осуществление некоторых условий, при наличии которых может наступить некоторое интересующее нас событие.

Приступим теперь к рассмотрению одной из главных схем теории вероятностей, имеющей, помимо прикладного значения в разнообразных областях знания, большое значение и в самой теории вероятностей, как математической науке. Эта схема состоит в том, что рассматривается последовательность взаимно независимых испытаний, т. е. таких испытаний, что вероятность того или иного результата в каждом из них не зависит от того, какие результаты наступили или наступят в остальных. В каждом из этих испытаний может наступить (или не наступить) некоторое событие A с вероятностью p, не зависящей от номера испытания. Описанная схема получила название *схемы Бернулли*: ее изучение ведет свое начало от известного швейцарского ученого Якова Бернулли, жившего в конце XVII века.

Мы уже имели дело со схемой Бернулли на примерах; чтобы убедиться в этом, достаточно вспомнить примеры предыдущего параграфа. Теперь же мы решим следующую общую задачу, частными случаями которой были все примеры, рассмотренные нами до сих пор в этой главе.

З а д а ч а. При некоторых условиях вероятность появления события A в каждом испытании равна p; найти вероятность того, что серия из n независимых испытаний даст k появлений и $n-k$ непоявлений события A.

Событие, вероятность которого надо найти, распадается на ряд разновидностей; для получения одной определенной разновидности мы должны произвольным образом выбрать из данной серии какие-нибудь k испытаний и допустить, что именно при этих k испытаниях произошло событие A, а при остальных $n-k$ оно не произошло. Таким образом, каждая такая разновидность требует наступления n определенных результатов, в том числе k появлений и $n-k$ непоявлений события A; мы получаем по правилу умножения, что вероятность каждой определенной разновидности равна

$$p^k(1-p)^{n-k};$$

число различных возможных разновидностей равно числу различных групп по k испытаний в каждой, которые можно составить из n испытаний, т. е. равно C_n^k Применяя правило сложения и известную формулу для числа сочетаний

$$C_n^k=\frac{n(n-1)\,\ldots\,(n-(k-1))}{k(k-1)\,\ldots\,2\cdot 1},$$

мы находим, что искомая вероятность k появлений события A при n независимых испытаниях равна

$$P_n(k)=\frac{n(n-1)\,\ldots\,(n-(k-1))}{k(k-1)\,\ldots\,2\cdot 1}\,p^k(1-p)^{n-k}, \qquad (3)$$

что и решает поставленную задачу. Часто бывает более удобно представлять выражение C_n^k в несколько ином виде; умножая числитель и знаменатель на произведение $(n-k)(n-(k+1))\ \ldots\ 2\cdot 1$, мы получаем:

$$C_n^k=\frac{n(n-1)\ \ldots\ 2\cdot 1}{k(k-1)\ \ldots\ 2\cdot 1\cdot(n-k)(n-(k+1))\ \ldots\ 2\cdot 1},$$

или, обозначая для краткости через $m!$ произведение всех целых чисел от 1 до m включительно,

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

для $P_n(k)$ это дает:

$$P_n(k)=\frac{n!}{k!\,(n-k)!}p^k(1-p)^{n-k}. \tag{4}$$

Формулы (3) и (4) обычно называют *формулами Бернулли.* При больших значениях n и k вычисление $P_n(k)$ по этим формулам затруднительно, так как факториалы $n!$, $k!$, $(n-k)!$ — очень большие и громоздко вычисляемые числа. Поэтому при вычислениях этого рода широко пользуются как специально составленными для факториалов таблицами, так и некоторыми приближенными формулами.

Пример. Вероятность того, что расход воды на некотором предприятии окажется нормальным (не больше определенного числа литров в сутки), равна 3/4. Найти вероятности того, что в ближайшие 6 дней расход воды будет нормальным в течение 1, 2, 3, 4, 5, 6 дней.

Обозначив через $P_6(k)$ вероятность того, что в течение k дней из шести расход воды будет нормальным, найдем по формуле (3) (где надо положить $p=3/4$)

$$P_6(6)=\left(\frac{3}{4}\right)^6=\frac{3^6}{4^6},$$

$$P_6(5)=6\cdot\left(\frac{3}{4}\right)^5\cdot\frac{1}{4}=\frac{6\cdot 3^5}{4^6},$$

$$P_6(4)=C_6^4\cdot\left(\frac{3}{4}\right)^4\cdot\left(\frac{1}{4}\right)^2=C_6^2\frac{3^4}{4^6}=\frac{6\cdot 5}{2\cdot 1}\cdot\frac{3^4}{4^6}=\frac{15\cdot 3^4}{4^6},$$

$$P_6(3) = C_6^3\left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^3 = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} \cdot \frac{3^3}{4^6} = \frac{20 \cdot 3^3}{4^6},$$

$$P_6(2) = \frac{6 \cdot 5}{2 \cdot 1}\left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^4 = \frac{15 \cdot 3^2}{4^6},$$

$$P_6(1) = 6 \cdot \frac{3}{4} \cdot \left(\frac{1}{4}\right)^5 = \frac{6 \cdot 3}{4^6};$$

наконец, очевидно, $P_6(0)$ (вероятность каждый день из 6 иметь перерасход) составляет $1/4^6$. Все шесть вероятностей выражаются дробями с одним и тем же знаменателем $4^6 = 4096$; этим мы, конечно, пользуемся для сокращения расчетов. Вычисления дают:

$$P_6(6) \approx 0{,}18; \quad P_6(5) \approx 0{,}36; \quad P_6(4) \approx 0{,}30;$$

$$P_6(3) \approx 0{,}13; \quad P_6(2) \approx 0{,}03; \quad P_6(1) \approx P_6(0) \approx 0.$$

Мы видим, что наиболее вероятен перерасход воды в течение одного или двух дней из шести и что вероятность перерасхода в течение пяти или шести дней $[P_6(1) + P_6(0)]$ практически равна нулю.

§ 15. Наивероятнейшее число наступлений события

Пример, который мы только что рассмотрели, показывает, что вероятность нормального расхода воды в течение ровно k дней с ростом числа k сначала возрастает, а затем, достигнув своего наибольшего значения, начинает убывать; яснее всего это видно, если изменение вероятности $P_6(k)$ с ростом числа k изобразить геометрически, в виде диаграммы, изображенной на рис. 4.

Еще более яркую картину дают диаграммы изменения величины $P_n(k)$ с ростом k, когда число n становится больше; так, при $n = 15$ и $p = 1/2$ диаграмма имеет вид, изображеный на рис. 5.

Для практики иногда требуется знать, какое число наступлений события является *наивероятнейшим*, т. е. при каком числе k вероятность $P_n(k)$ наибольшая (при этом, конечно, p и n предполагаются заданными).

Формулы Бернулли позволяют во всех случаях найти простое решение поставленного вопроса; этим мы теперь и займемся.

Вычислим сначала величину отношения $\frac{P_n(k+1)}{P_n(k)}$. В силу формулы (4),

$$P_n(k+1)=\frac{n!}{(k+1)!\,(n-k-1)!}p^{k+1}(1-p)^{n-k-1}, \qquad (5)$$

а из формул (3) и (5) получаем:

$$\frac{P_n(k+1)}{P_n(k)}=\frac{n!k!\,(n-k)!\,p^{k+1}(1-p)^{n-k-1}}{(k+1)!\,(n-k-1)!\,n!\,p^k(1-p)^{n-k}}=\frac{n-k}{k+1}\,\frac{p}{1-p}.$$

Вероятность $P_n(k+1)$ будет больше, равна или меньше вероятности $P_n(k)$, смотря по тому, будет ли отношение $\frac{P_n(k+1)}{P_n(k)}$ больше, равно или меньше единицы, а это, как мы видим, сводится к вопросу о том, какое из трех соотношений

$$\frac{n-k}{k+1}\,\frac{p}{1-p}>1, \quad \frac{n-k}{k+1}\cdot\frac{p}{1-p}=1, \quad \frac{n-k}{k+1}\cdot\frac{p}{1-p}<1 \qquad (6)$$

окажется верным. Если мы, например, хотим выяснить, при каких значениях k выполняется $P_n(k+1) > P_n(k)$, то для этого нужно узнать, при каких значениях k имеет место неравенство

$$\frac{n-k}{k+1}\,\frac{p}{1-p}>1, \text{ или } (n-k)\,p>(k+1)(1-p),$$

откуда получаем:

$$np-(1-p)>k;$$

таким образом, покуда k, возрастая, не достигнет величины $np-(1-p)$, мы будем все время иметь $P_n(k+1) > P_n(k)$, т. е. с ростом числа k вероятность $P_n(k)$ будет все время возрастать. Так, например, в схеме, которой соответствует диаграмма на рис. 5, $p=1/2$, $n=15$, $np-(1-p)=7$; значит покуда $k<7$, т. е. для всех k, от 0 до 6 включительно, мы имеем $P_n(k+1) > P_n(k)$. Это и подтверждает диаграмма.

Совершенно подобным же образом, исходя из двух других соотношений (6), мы находим, что

$$P_n(k+1)=P_n(k), \quad \text{если} \quad k=np-(1-p),$$

$$P_n(k+1)<P_n(k), \quad \text{если} \quad k>np-(1-p);$$

таким образом, как только число k, возрастая, перешагнет через грань $np-(1-p)$, так вероятность $P_n(k)$ начнет убывать и будет убывать до $P_n(n)$.

Этот вывод прежде всего убеждает нас, что усмотренное нами выше на примерах поведение величины $P_n(k)$ при росте числа k (сперва возрастание, потом убывание) является общим законом, имеющим место во всех случаях. Но более того — этот вывод позволяет

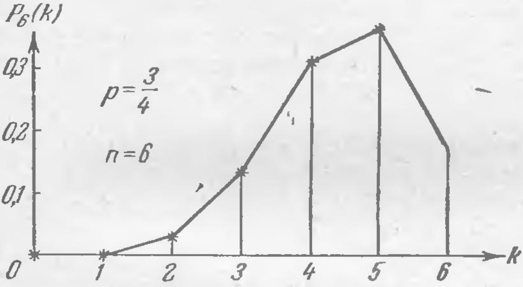

Рис. 4.

нам также решить немедленно ту задачу, которую мы себе поставили, — определить наивероятнейшее значение числа k. Обозначим это наивероятнейшее значение через k_0. Тогда

$$P_n(k_0+1)\leqslant P_n(k_0),$$

откуда, по предыдущему,

$$k_0\geqslant np-(1-p);$$

с другой стороны,

$$P_n(k_0-1)\leqslant P_n(k_0),$$

для чего, по предыдущему, должно быть

$$k_0-1\leqslant np-(1-p)$$

или

$$k_0\leqslant np-(1-p)+1=np+p.$$

Таким образом, наивероятнейшее значение k_0 числа k должно удовлетворять двойному неравенству

$$np-(1-p)\leqslant k_0\leqslant np+p; \tag{7}$$

промежуток от $np-(1-p)$ до $np+p$, в котором должно, таким образом, лежать число k_0, имеет величину 1, как показывает простое вычитание; поэтому если

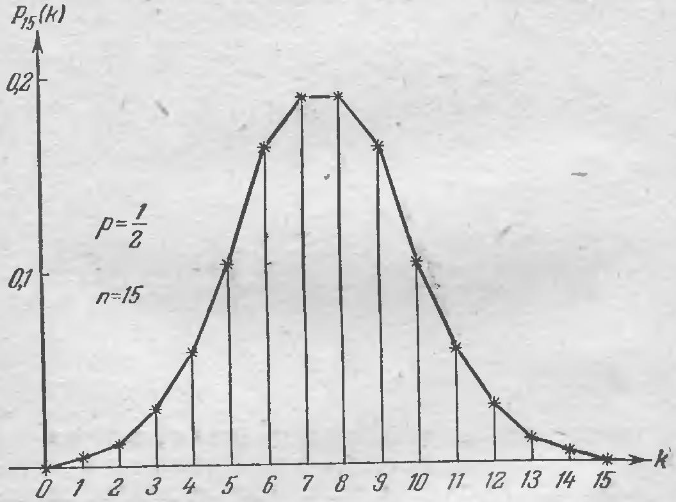

Рис. 5.

какой-либо из концов этого промежутка, например число $np-(1-p)$, не есть целое число, то между этими концами будет обязательно лежать одно и только одно целое число и k_0 будет однозначно определено. Этот случай мы должны рассматривать как нормальный: ведь $p<1$, и потому лишь в исключительных случаях величина $np-(1-p)$ будет целым числом. В этом исключительном случае неравенства (7) дают для числа k_0 два значения: $np-(1-p)$ и $np+p$, отличающиеся друг от друга на единицу; эти два значения и будут наивероятнейшими; их вероятности равны между собою и превышают вероятности всех других значений числа k. Именно этот исключительный случай

чай имеет место, например, в схеме, изображенной диаграммой на рис. 5; здесь $n = 15$, $p = 1/2$ и, значит, $np - (1 - p) = 7$, $np + p = 8$; наиболее вероятными значениями числа k наступлений события служат числа 7 и 8; их вероятности равны между собою, каждая из них приближенно равна 0,196 (все это можно видеть на диаграмме).

Пример 1. В результате многолетних наблюдений для некоторой местности было выяснено, что вероятность того, что в течение 1 июля выпадет дождь, равна 4/17. Найти наивероятнейшее число дождливых дней 1 июля за ближайшие 50 лет. Здесь $n = 50$, $p = 4/17$,

$$np - (1 - p) = 50 \cdot \frac{4}{17} - \frac{13}{17} = 11;$$

число это оказалось целым; значит, мы имеем дело с исключительным случаем; наивероятнейшим значением числа дождливых дней будут равновероятные между собой числа 11 и 12.

Пример 2. В одном физическом эксперименте производятся наблюдения за частицами определенного типа. При одних условиях за промежуток времени определенной длины в среднем появляется 60 частиц и каждая из них с вероятностью 0,7 имеет скорость большую, чем v_0. При других условиях за тот же промежуток времени в среднем появляется лишь 50 частиц, но для каждой из них вероятность имеет скорость, превышающую v_0, равна 0,8. Для каких условий опыта вероятнейшее число частиц со скоростью, превосходящей v_0, больше?

Для первых условий эксперимента $n = 60$, $p = 0,7$, $np - (1 - p) = 41,7$, $k_0 = 42$.

Для вторых условий эксперимента $n = 50$, $p = 0,8$, $np - (1 - p) = 39,8$, $k_0 = 40$.

Мы видим, что вероятнейшее число «быстрых» частиц в первых условиях эксперимента несколько больше.

В практике часто встречаются положения, когда число n весьма велико (массовая стрельба, массовое производство изделий и т. п.). В этом случае и произведение np будет очень большим числом (если только вероятность p не чрезвычайно мала). А так как в выражениях $np-(1-p)$ и $np+p$, между которыми заключено наивероятнейшее число появлений события, вторые члены (т. е. p и $1-p$) меньше единицы, то оба эти числа, а значит и заключенное между ними наивероятнейшее число появлений события, близки к np. Так, если вероятность соединения абонентов за срок, меньший 15 сек., равна 0,74, то мы можем принять $1000 \cdot 0{,}74$ за вероятнейшее число соединений абонентов за срок, меньший 15 сек., из каждой тысячи вызовов, поступающих на телефонную станцию.

Этому выводу можно придать еще более точную форму. Если k_0 означает наивероятнейшее число появлений события при n испытаниях, то k_0/n есть наивероятнейшая «доля» появлений события при тех же n испытаниях; неравенства (7) дают нам:

$$p-\frac{1-p}{n} \leqslant \frac{k_0}{n} \leqslant p+\frac{p}{n}.$$

Представим себе, что мы, оставляя вероятность p появления события для отдельного испытания неизменной, будем увеличивать все более и более число испытаний n (при этом, конечно, будет увеличиваться и наивероятнейшее число появлений k_0). Дроби $(1-p)/n$ и p/n, стоящие в левой и правой частях последних неравенств, будет становиться все меньше и меньше; значит, при больших n этими дробями можно пренебречь, т. е. можно считать левую и правую части неравенств, а значит и дробь k_0/n, равными p. Таким образом, *наивероятнейшая доля появлений события при большом числе испытаний практически равна вероятности появления события при отдельном испытании.*

Так, если при некоторых измерениях вероятность совершить в отдельном измерении ошибку, заключенную между α и β, равна 0,84, то при большом числе измерений с наибольшей вероятностью можно ожидать примерно в 84% случаев ошибок, заключенных между α и β. Это не значит, конечно, что вероятность получить ровно 84% таких ошибок будет велика; напротив, сама эта «наибольшая вероятность» при большом числе измерений будет очень малой (так, мы видели в схеме рис. 5, что наибольшая вероятность оказалась равной 0,196, причем там речь шла всего о 15 испытаниях; при большем числе испытаний она значительно меньше). Эта вероятность является наибольшей только в сравнительном смысле: вероятность получить 84% измерений с ошибками, заключенными между α и β, больше, чем вероятность получить 83% или 86% таких измерений.

С другой стороны, легко понять, что при длинных сериях измерений вероятность того или иного отдельного числа ошибок данной величины не может представлять значительного интереса. Если мы, например, производим 200 измерений, то вряд ли целесообразно вычислять вероятность того, что ровно 137 из них будут измерены с заданной точностью, так как практически безразлично, будет это число равно 137, 136 или 138 и даже хотя бы 140. Напротив, вопрос, например, о вероятности того, что число измерений, для которых ошибка заключена в данных границах, будет более 100 из 200 произведенных измерений, или что это число будет заключено между 100 и 125, или что оно окажется меньшим 50 и т. д., представляет несомненный интерес для практики. Как выразить такого рода вероятности? Пусть мы хотим, например, найти вероятность того, что число измерений будет заключено между 100 и 120 (включая 120); точнее, будем искать вероятность неравенств

$$100 < k \leqslant 120,$$

где k — число попаданий. Для того чтобы осуществились эти неравенства, нужно, чтобы k оказалось равным одному из двадцати чисел 101, 102, ..., 120; по правилу сложения эта вероятность равна

$$\mathrm{P}(100 < k \leqslant 120) = \\ = P_{200}(101) + P_{200}(102) + \ldots + P_{200}(120);$$

для непосредственного вычисления этой суммы нам пришлось бы предварительно вычислить 20 отдельных вероятностей типа $P_n(k)$ по формуле (3); *при столь* больших числах подобные вычисления представляют непреодолимые трудности; поэтому суммы полученного нами вида на практике никогда не вычисляют непосредственно. Для этого существуют удобные приближенные формулы и таблицы. Составление этих формул и таблиц основано на сложных приемах математического анализа, которых мы здесь касаться не будем, однако можно простыми рассуждениями и о вероятностях типа $\mathsf{P}(100 < k \leqslant 120)$ получить сведения, во многих случаях приводящие к исчерпывающему решению поставленных задач. Об этом мы будем говорить в следующей главе.

ГЛАВА ШЕСТАЯ

ТЕОРЕМА БЕРНУЛЛИ

§ 16. Содержание теоремы Бернулли

Вглядимся еще раз в диаграмму на рис. 5 (стр. 58), где вероятности различных значений числа k — появлений рассматриваемого события — числа $P_{15}(k)$ изображены вертикальными черточками. Вероятность, приходящаяся на какой-нибудь *участок* значений k, т. е. вероятность того, что число появлений интересующего нас события окажется равным какому-нибудь из чисел этого участка, по правилу сложения равна сумме вероятностей всех чисел этого участка, т. е. равна сумме длин всех вертикальных черточек, расположенных над этим участком. Чертеж наглядно показывает, что эта сумма для различных участков одной и той же длины весьма различна. Так, участки $2 \leqslant k < 5$ и $7 \leqslant k < 10$ имеют одинаковую длину; вероятность каждого из них выражается суммой длин трех вертикальных черточек, и мы видим, что для второго участка сумма эта значительно больше, чем для первого. Мы уже знаем, что диаграммы вероятностей $P_n(k)$ для всех n имеют такой же, в основном, вид, что и диаграмма на рис. 5, т. е. величина $P_n(k)$ с ростом k сначала возрастает, а потом, после прохождения через свое наибольшее значение, убывает; поэтому ясно, что из двух участков значений числа k, имеющих одинаковую длину, во всех случаях бóльшую вероятность будет иметь тот, который расположен ближе к наивероятнейшему значению k_0. В частности, на участок, имеющий серединой число k_0, всегда будет приходиться бо́льшая вероятность, чем на любой другой участок той же длины.

Но оказывается, что в этом направлении можно высказать гораздо больше. Всех возможных значений числа k появлений события при n испытаниях имеется $n+1$ $(0 \leqslant k \leqslant n)$.

Возьмем участок, имеющий середину в k_0 и заключающий в себе только малую долю, например одну

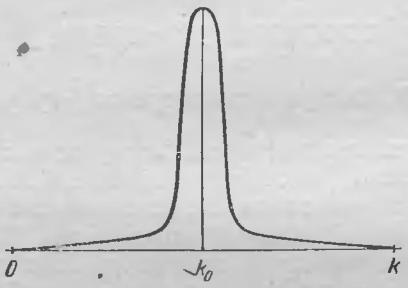

Рис. 6.

сотую, возможных значений числа k; тогда оказывается, что если общее число испытаний n очень большое, то на этот участок приходится подавляющая вероятность, а все другие вместе взятые значения числа k имеют вероятность ничтожно малую. Таким образом, хотя выбранный нами участок и ничтожно мал по сравнению с n (на чертеже он займет всего одну сотую долю всей длины диаграммы), но тем не менее сумма расположенных над ним вертикальных черточек будет значительно больше, чем сумма всех остальных вертикальных черточек. Причина этого в том, что черточки центральной части диаграммы во много раз крупнее черточек, располагающихся по краям. Таким образом, для больших n диаграмма величины $P_n(k)$ имеет вид, примерно изображенный на рис. 6.

Практически все это, очевидно, означает следующее: *если мы производим серию из большого числа n испытаний, то с вероятностью, близкой к единице, мы можем ожидать, что число k появлений события A бу-*

дет очень близко к своему наивероятнейшему значению, отличаясь от него лишь на незначительную долю общего числа n произведенных испытаний.

Это предложение, известное под именем *теоремы Бернулли* и открытое в начале восемнадцатого столетия, представляет собой один из важнейших законов теории вероятностей: до середины прошлого столетия все доказательства этой теоремы требовали сложных математических средств, и только великий русский математик П. Л. Чебышев впервые нашел очень простое и краткое обоснование этого закона; замечательное доказательство Чебышева мы теперь и приведем.

§ 17. Доказательство теоремы Бернулли

Мы уже знаем, что при большом числе испытаний n наивероятнейшее число k_0 появлений события A почти не отличается от величины np, где p, как всегда,

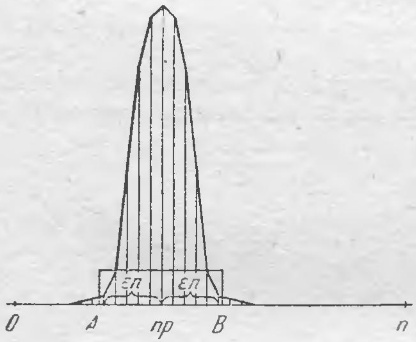

Рис. 7.

означает вероятность события A для отдельного испытания. Поэтому нам достаточно доказать, что при большом числе испытаний число k появлений события A будет с подавляющей вероятностью отличаться от np весьма мало — не более чем на сколь угодно малую долю числа n (не более чем, например, на $0{,}01\,n$ или $0{,}001\,n$, или вообще не более чем на εn, где ε — как угодно малое число); иначе говоря, мы должны показать, что вероятность

$$\mathsf{P}(|k-np|>\varepsilon n) \tag{1}$$

для достаточно больших n будет как угодно малой.

Чтобы в этом убедиться, заметим, что по правилу сложения вероятность (1) равна сумме вероятностей $P_n(k)$ для всех тех значений числа k, которые отстоят от np больше чем на εn; на нашей обычной диаграмме (рис. 7) эта сумма изображается суммою длин всех вертикальных черточек, лежащих вне отрезка AB, как вправо, так и влево от него. Так как общая сумма всех вертикальных черточек (как сумма вероятностей полной системы событий) равна единице, то это значит, что подавляющая, почти равная единице часть этой суммы приходится на долю отрезка AB и лишь ничтожно малая часть ее — на области, лежащие вне этого отрезка.

Итак,

$$\mathsf{P}(|k-np|>\varepsilon n)=\sum_{|k-np|>\varepsilon n} P_n(k). \tag{2}$$

Обращаемся теперь к рассуждению Чебышева. Так как в каждом члене написанной суммы

$$\left|\frac{k-np}{\varepsilon n}\right|>1,$$

а значит, и

$$\left(\frac{k-np}{\varepsilon n}\right)^2>1,$$

то мы можем только увеличить эту сумму, если каждый ее член $P_n(k)$ заменим выражением

$$\left(\frac{k-np}{\varepsilon n}\right)^2 P_n(k);$$

поэтому

$$\mathsf{P}(|k-np|>\varepsilon n)<\sum_{|k-np|>\varepsilon n}\left(\frac{k-np}{\varepsilon n}\right)^2 P_n(k)=$$

$$=\frac{1}{\varepsilon^2 n^2}\sum_{|k-np|>\varepsilon n}(k-np)^2 P_n(k);$$

далее, очевидно, что последнюю сумму мы увеличим еще больше, если в дополнение к уже имеющимся в ней слагаемым добавим еще новые, заставляя число k пробегать не только участки до $np-\varepsilon n$ и от $np+\varepsilon n$, но и весь ряд возможных для него значений, т. е. весь ряд чисел от 0 до n включительно; мы получаем таким образом, что и подавно

$$\mathsf{P}(|k-np|>\varepsilon n)<\frac{1}{\varepsilon^2 n^2}\sum_{k=0}^{n}(k-np)^2 P_n(k). \qquad (3)$$

Последняя сумма выгодно отличается от всех предыдущих тем, что ее можно в точности вычислить; метод Чебышева и заключается в замене трудно оцениваемой суммы (2) суммой (3), допускающей точное вычисление.

Мы приступаем теперь к этому вычислению; как бы длинным оно нам ни показалось, это уже — трудности технического порядка, с которыми справится всякий, кто знает алгебру; замечательная же идея Чебышева нами уже вполне использована, ибо она состоит именно в переходе от равенства (2) к неравенству (3).

Прежде всего мы легко находим:

$$\sum_{k=0}^{n}(k-np)^2 P_n(k)=$$

$$=\sum_{k=0}^{n}k^2P_n(k)-2np\sum_{k=0}^{n}kP_n(k)+n^2p^2\sum_{k=0}^{n}P_n(k). \qquad (4)$$

Из трех сумм правой части последняя равна единице, как сумма вероятностей полной системы событий. Значит, нам остается только вычислить суммы

$$\sum_{k=0}^{n}kP_n(k) \quad \text{и} \quad \sum_{k=0}^{n}k^2P_n(k);$$

при этом в обеих суммах члены, соответствующие $k=0$, равны нулю, так что можно начинать суммирование с $k=1$.

5*

1) Для вычисления обеих сумм выразим $P_n(k)$ по формуле (4) главы 5 (стр. 54). Имеем

$$\sum_{k=1}^{n} kP_n(k) = \sum_{k=1}^{n} \frac{kn!}{k!\,(n-k)!} p^k (1-p)^{n-k};$$

так как, очевидно, $n! = n(n-1)!$ и $k! = (k-1)!$, то получаем:

$$\sum_{k=1}^{n} kP_n(k) =$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,[(n-1)-(k-1)]!} p^{k-1} (1-p)^{(n-1)-(k-1)},$$

или, полагая в сумме правой части $k-1=l$ и замечая, что l изменяется от 0 до $n-1$ при изменении k от 1 до n,

$$\sum_{k=1}^{n} kP_n(k) = np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!\,(n-1-l)!} p^l (1-p)^{n-1-l} =$$

$$= np \sum_{l=0}^{n-1} P_{n-1}(l).$$

Последняя сумма $\sum_{l=0}^{n-1} P_{n-1}(l)$, разумеется, равна единице, ибо эта есть сумма вероятностей полной системы событий — всевозможных чисел l появлений события при $n-1$ испытаниях. Таким образом, для суммы $\sum_{k=0}^{n} kP_n(k)$ получили чрезвычайно простое выражение:

$$\sum_{k=0}^{n} kP_n(k) = np. \qquad (5)$$

2) Для вычисления второй суммы найдем сначала величину $\sum_{k=1}^{n} k(k-1)P_n(k)$; так как член, соответствующий $k=1$, очевидно, равен нулю, то суммирование

можно начать со значения $k=2$. Замечая, что $n! = n(n-1)(n-2)!$ и что $k! = k(k-1)(k-2)!$, мы легко заключаем, полагая подобно предыдущему $k-2=m$:

$$\sum_{k=1}^{n} k(k-1)P_n(k) = \sum_{k=2}^{n} k(k-1)P_n(k) =$$

$$= \sum_{k=2}^{n} \frac{k(k-1)\,n!}{k!\,(n-k)!} p^k (1-p)^{n-k} =$$

$$= n(n-1)p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!\,[(n-2)-(k-2)]!} \times$$

$$\times p^{k-2}(1-p)^{(n-2)-(k-2)} =$$

$$= n(n-1)p^2 \sum_{m=0}^{n-2} \frac{(n-2)!}{m!\,(n-2-m)!} p^m (1-p)^{n-2-m} =$$

$$= n(n-1)p^2 \sum_{m=0}^{n-2} P_{n-2}(m) = n(n-1)p^2, \tag{6}$$

ибо последняя сумма снова равна единице, как сумма вероятностей некоторой полной системы событий — всевозможных чисел появлений события при $n-2$ испытаниях.

Наконец, формулы (5) и (6) дают:

$$\sum_{k=1}^{n} k^2 P_n(k) = \sum_{k=1}^{n} k(k-1)P_n(k) + \sum_{k=1}^{n} kP_n(k) =$$
$$= n(n-1)p^2 + np = n^2p^2 + np(1-p). \tag{7}$$

Теперь обе нужные нам суммы вычислены. Вставляя результаты (5) и (7) в соотношение (4), окончательно находим:

$$\sum_{k=0}^{n} (k-np)^2 P_n(k) = n^2p^2 + np(1-p) - 2np \cdot np + n^2p^2 =$$
$$= np(1-p);$$

вставляя же полученное чрезвычайно простое выражение в неравенство (3), получаем:

$$\mathrm{P}(|k-np|>\varepsilon n)<\frac{np(1-p)}{\varepsilon^2 n^2}=\frac{p(1-p)}{\varepsilon^2 n}. \qquad (8)$$

Это неравенство доказывает все необходимое. В самом деле, число ε можно взять, правда, как угодно малым, но, выбрав, мы уже больше его не меняем; число же испытаний n, по смыслу нашего утверждения, может быть как угодно большим; поэтому дробь $\frac{p(1-p)}{\varepsilon^2 n}$ можно предположить как угодно малой, так как с возрастанием n знаменатель ее может быть как угодно большим, а числитель при этом не меняется.

Пусть, например, $p=0{,}75$, так что $1-p=0{,}25$, и

$$p(1-p)=0{,}1875<0{,}2;$$

выберем $\varepsilon=0{,}01$; тогда неравенство (8) дает

$$\mathrm{P}\left(\left|k-\frac{3}{4}n\right|>\frac{1}{100}n\right)<\frac{0{,}2}{0{,}0001n}=\frac{2000}{n}.$$

Если, например, $n=200\,000$, то

$$\mathrm{P}(|k-150\,000|>2000)<0{,}01.$$

Практически это означает, например, следующее: если на некотором производстве при установившемся технологическом процессе в среднем 75% изделий обладают некоторым свойством (например, принадлежат к 1-му сорту), то из 200 000 изделий с вероятностью, превышающей 0,99 (т. е. практически почти достоверно), этим свойством будут обладать от 148 000 до 152 000 изделий.

К этому необходимо сделать два замечания:

1. Неравенство (8) дает весьма грубую оценку для вероятности $\mathrm{P}(|k-np|>\varepsilon n)$; на самом деле эта вероятность, особенно при больших значениях n, значительно меньше; на практике пользуются поэтому более *точными оценками*, обоснование которых, однако, значительно сложнее.

2. Оценка, даваемая неравенством (8), становится значительно более точной, когда вероятность p очень мала или, напротив, очень близка к единице.

Так, если в только что приведенном примере вероятность того, что изделие обладает некоторым свойством, равна $p = 0{,}95$, то

$$1 - p = 0{,}05, \quad p(1-p) < 0{,}05.$$

Поэтому, выбирая $\varepsilon = 0{,}005$, $n = 200\,000$, находим:

$$\frac{p(1-p)}{\varepsilon^2 n} < \frac{0{,}05 \cdot 1\,000\,000}{25 \cdot 200\,000} = 0{,}01,$$

как и прежде. Но теперь εn равно не 2000, а только 1000; отсюда (так как $np = 190\,000$) заключаем, что с практической достоверностью число изделий, обладающих рассматриваемым свойством, при общем числе 200 000 изделий окажется заключенным между 189 000 и 191 000.

Таким образом, при $p = 0{,}95$ неравенство (8) практически гарантирует, что для ожидаемого числа изделий с интересующим нас свойством промежуток имеет вдвое меньшую длину, чем при $p = 0{,}75$, ибо

$$\mathsf{P}(|k - 190\,000| > 1000) < 0{,}01.$$

З а д а ч а. Известно, что одна четвертая часть рабочих некоторой отрасли промышленности имеет среднее образование. Для некоторого обследования наудачу выбрано 200 000 рабочих. Найти 1) наивероятнейшее значение числа рабочих со средним образованием среди выбранных 200 000 и 2) вероятность того, что истинное (фактическое) число таких рабочих отклонится от наивероятнейшего не более чем на 1,6%.

При решении задачи мы исходим из того факта, что вероятность иметь среднее образование равна одной четверти для каждого из наудачу выбираемых 200 000 рабочих (в этом и заключается смысл слова «наудачу»). Таким образом,

$$n = 200\,000, \quad p = \frac{1}{4}, \quad k_0 = np = 50\,000; \quad p(1-p) = \frac{3}{16}.$$

Надо найти вероятность того, что $|k - np| < 0{,}016\,np$

или $|k - np| < 800$, где k — число рабочих со средним образованием.

Выбираем ε так, чтобы иметь $\varepsilon n = 800$; отсюда получаем: $\varepsilon = \frac{800}{n} = 0{,}004$. Формула (8) дает:

$$P(|k - 50\,000| > 800) < \frac{3}{16 \cdot 0{,}000016 \cdot 200\,000} \approx 0{,}06,$$

откуда

$$P(|k - 50\,000| < 800) > 0{,}94.$$

Ответ. Искомое наивероятнейшее значение равно 50 000, искомая вероятность больше чем 0,94.

На самом деле искомая вероятность значительно ближе к единице.

ЧАСТЬ ВТОРАЯ
СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

ГЛАВА СЕДЬМАЯ
СЛУЧАЙНАЯ ВЕЛИЧИНА И ЗАКОН РАСПРЕДЕЛЕНИЯ

§ 18. Понятие случайной величины

Во всем предыдущем мы много раз уже встречались с такими величинами, численное значение которых не может быть раз навсегда определено, а меняется под влиянием случайных воздействий. Так, число мальчиков на сотню новорожденных не будет для всех сотен одним и тем же. Или длина волокон хлопка определенного сорта меняется весьма сильно не только для различных районов, где этот сорт возделывается, но даже для одного куста и одной коробочки.

Приведем еще несколько примеров величин такого рода.

1) Стреляя из одного и того же орудия на одном и том же прицеле и при неизменных условиях, мы тем не менее наблюдаем, что снаряды ложатся в разных местах; это явление называют «рассеиванием» снарядов. Расстояние места падения снаряда от места его вылета есть величина, при различных выстрелах получающая, в зависимости от не поддающихся предварительному учету *случайных* обстоятельств, различные численные значения.

2) Скорость молекулы газа не остается неизменной, а меняется от столкновений с другими молекулами. Ввиду того что каждая молекула может либо столкнуться, либо не столкнуться с каждой другой молекулой газа, изменение ее скорости носит чисто *случайный* характер.

3) Число метеоритов, падающих на землю в течение года, достигающих ее поверхности, не постоянно, а подвержено значительным колебаниям в зависимости от целого ряда обстоятельств случайного характера.

4) Вес зерен пшеницы, выращенных на некотором участке не равен некоторой определенной величине, а меняется от одного зерна к другому. В силу невозможности учесть влияние всех факторов (качество участка почвы, на котором вырос колос с данным зерном, условия освещения зерна, водный режим и пр.), определяющих рост зерна, его вес является величиной, меняющейся в зависимости от случая.

Несмотря на всю разнородность рассмотренных примеров, все они с интересующей нас теперь точки зрения представляют собой одну и ту же картину. В каждом из этих примеров мы имеем дело с величиной, так или иначе характеризующей собой результат предпринятой операции (например, счета метеоритов, измерения длин волокон); каждая из этих величин при различных операциях, какими бы однородными мы ни старались сделать условия их осуществления, может принимать различные значения, в зависимости от ускользающих от нашего контроля случайных различий в обстановке этих операций. Такого рода величины мы называем в теории вероятностей *случайными величинами*; приведенных нами примеров уже достаточно, чтобы убедить нас, сколь важным может быть изучение случайных величин для приложений теории вероятностей к самым разнообразным областям знания и практики.

Знать данную случайную величину — это не значит, конечно, знать ее численное значение, ибо если мы, например, знаем, что конденсатор до пробоя проработал 5324 часа, то тем самым уже время его бесперебойной работы приняло определенное численное значение и перестало быть случайной величиной. Что же мы должны знать о случайной величине, для того чтобы иметь всю возможную полноту сведений о ней именно как о случайной величине? Очевидно, что для этого прежде всего мы должны знать все численные значения, которые она способна принимать. Так, если при испытании электронных ламп на срок службы минимальное время оказалось 2306 часов, а максимальное — 12 108, то срок службы лампы может принимать все значения, заключенные между этими двумя границами. В примере 3) очевидно, что число метеоритов, достигших земной поверхности в течение года, может иметь своим численным значением любое целое неотрицательное число, т. е. 0, 1, 2, 3 и т. д.

Однако знание лишь одного перечня возможных значений случайной величины не дает о ней еще таких сведений, которые могли бы служить материалом для практически необходимых оценок. Так, если во втором примере рассмотреть газ при двух разных температурах, то возможные численные значения скоростей молекул для них одни и те же, и, следовательно, перечень этих значений не дает возможности какой-либо сравнительной оценки этих температур.

Между тем, различие температур указывает на весьма существенное различие в состоянии газа — различие, о котором один только перечень возможных значений скоростей молекул не дает нам никаких представлений. Если мы хотим оценить температуру данной массы газа, а нам дают только список возможных значений скоростей его молекул, то мы, естественно, спрашиваем, как часто наблюдается та или другая скорость. Другими словами, мы, естественно, стремимся узнать *вероятности* различных возможных значений интересующей нас случайной величины.

§ 19. Понятие закона распределения

Возьмем для начала совсем простой пример. Производится стрельба по мишени, изображенной на рис. 8; попадание в область I дает стреляющему три очка, в область II — два очка и в область III — одно очко *).

*) Читатель может возразить, сказав, что за попадание в область III, т. е. за промах, не следует давать и одного очка. Однако, если очко является правом на выстрел, то тот, кто произвел плохой выстрел, тем самым уже получил одно очко.

Рассмотрим в качестве случайной величины число очков, выбиваемых при отдельном выстреле; возможными значениями здесь служат числа 1, 2 и 3; обозначим соответственно через p_1, p_2 и p_3 вероятности этих трех значений, так что, например, p_3 означает вероятность попадания в область I мишени. В то время как возможные значения нашей случайной величины для всех стрелков одни и те же, вероятности p_1, p_2 и p_3 для различных стрелков могут весьма существенно различаться друг от друга, и очевидно, что этим различием и определяется различие в качестве стрельбы. Так, для очень хорошего стрелка может быть, например, что $p_3 = 0{,}8$, $p_2 = 0{,}2$; $p_1 = 0$; для среднего стрелка $p_3 = 0{,}3$; $p_2 = 0{,}5$; $p_1 = 0{,}2$; для совсем плохого стрелка $p_3 = 0{,}1$; $p_2 = 0{,}3$; $p_1 = 0{,}6$.

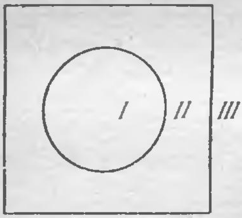

Рис. 8.

Если при некоторой стрельбе производится 12 выстрелов, то возможными значениями числа попаданий в каждую из областей I, II и III будут служить все целые числа от 0 до 12 включительно; но сам по себе этот факт еще не дает нам возможности судить о качестве стрельбы; напротив, мы можем составить себе исчерпывающее представление об этом качестве, если, кроме возможных значений числа попаданий, нам *даются и вероятности этих значений*, т. е. числа, показывающие, как часто среди серий по 12 выстрелов будет встречаться то или другое число попаданий в ту или иную область.

Ясно, что так будет обстоять дело и во всех случаях; зная вероятности различных возможных значений случайной величины, мы тем самым будем знать, как часто следует ожидать появления более выгодных или менее выгодных ее значений; а этого, очевидно, достаточно для суждения об эффективности или доброкачественности той операции, с которой связана данная случайная величина. Практика показывает, что знания вероятностей всех возможных значений изучаемой случайной величины действительно достаточно для решения любого вопроса, связанного с оценкой этой случайной величины как показателя доброкачественности соответствующей операции. Итак, мы приходим к выводу, что для полной характеристики той или другой случайной величины, как таковой, необходимо и достаточно знать:

1) перечень всех возможных значений этой случайной величины и

2) вероятность каждого из этих значений.

Отсюда видно, что задавать случайную величину удобно посредством таблицы из двух строк: верхняя строка содержит в каком-нибудь порядке возможные значения случайной величины, а нижняя — их вероятности, так что под каждым из возможных значений стоит его вероятность. Так, в рассмотренном выше примере число очков, выбиваемых при одном выстреле лучшим из стрелков, может, как случайная величина, быть представлено таблицей

1	2	3
0	0,2	0,8

(I)

В общем случае случайная величина, возможные значения которой суть $x_1, x_2, \ldots, x_n$, а соответствующие вероятности — $p_1, p_2, \ldots, p_n$, определяется таблицей

x_1	x_2	...	x_n
p_1	p_2	...	p_n

Задать такую таблицу, т. е. задать все возможные значения случайной величины вместе с их вероятностями, означает, как говорят, задать *закон распределения* этой случайной величины. Знание закона распределения данной случайной величины позволяет решать все связанные с нею вопросы.

Задача. Число очков, выбиваемых при одном выстреле одним стрелком, имеет закон распределения (I);

такое же число очков для другого стрелка имеет закон распределения

1	2	3
0,2	0,5	0,3

(II)

Найти закон распределения для суммы очков, выбиваемых обоими стрелками.

Ясно, что сумма, о которой идет речь, — случайная величина; наша задача — составить ее таблицу. Для этого мы должны рассмотреть все возможные результаты совместной стрельбы наших двух стрелков; мы расположим эти результаты в следующую таблицу, где вероятность каждого результата вычисляется по правилу умножения для независимых событий и где x означает число очков, выбиваемых первым стрелком, а y — число очков, выбиваемых вторым стрелком.

№ результата	x	y	$x+y$	Вероятность результата
1)	1	1	2	$0 \cdot 0{,}2 = 0$
2)	1	2	3	$0, \cdot 05 = 0$
3)	1	3	4	$0 \cdot 0{,}3 = 0$
4)	2	1	3	$0{,}2 \cdot 0{,}2 = 0{,}04$
5)	2	2	4	$0{,}2 \cdot 0{,}5 = 0{,}1$
6)	2	3	5	$0{,}2 \cdot 0{,}3 = 0{,}06$
7)	3	1	4	$0{,}8 \cdot 0{,}2 = 0{,}16$
8)	3	2	5	$0{,}8 \cdot 0{,}5 = 0{,}4$
9)	3	3	6	$0{,}8 \cdot 0{,}3 = 0{,}24$

Эта таблица показывает, что интересующая нас сумма $x + y$ может принимать значения 3, 4, 5 и 6;

значение 2 для нее невозможно, так как вероятность его равна нулю *). Мы имеем $x+y=3$ в случаях результатов 2) и 4); стало быть, для того чтобы сумма $x+y$ получила значение 3, надобно наступление одного из результатов 2) или 4); вероятность этого по правилу сложения равна сумме вероятностей этих результатов, т. е. равна $0+0{,}04=0{,}04$. Для равенства

$$x+y=4$$

надобно наступление одного из результатов 3), 5) или 7); вероятность этого равенства равна поэтому (опять по правилу сложения) $0+0{,}1+0{,}16=0{,}26$; подобным же образом мы находим, что вероятность значения 5 для суммы $x+y$ равна

$$0{,}06+0{,}4=0{,}46,$$

а вероятность значения 6, наступающего только в случае результата 9), равна 0,24. Таким образом, для случайной величины $x+y$ мы получаем следующую таблицу возможных значений:

3	4	5	6
0,04	0,25	0,46	0,24

(III)

Эта таблица (III) полностью решает поставленную задачу.

Сумма всех четырех вероятностей в таблице (III) равна единице; этим свойством должен, конечно, обладать каждый закон распределения, так как речь идет о сумме вероятностей всех возможных значений случайной величины, т. е. о сумме вероятностей некоторой полной системы событий. Этим свойством законов распределения удобно пользоваться как контрольным приемом для проверки правильности произведенных вычислений.

*) Можно, конечно, считать число 2 и возможным значением величины $x+y$, имеющим вероятность 0, подобно тому, как мы ради общности сделали это для значения 1 в таблице (I).

ГЛАВА ВОСЬМАЯ

СРЕДНИЕ ЗНАЧЕНИЯ

§ 20. Определение среднего значения случайной величины

Тє два стрелка, о которых мы только что говорили, стреляя вместе, могут выбить, в зависимости от случайных обстоятельств, либо 3, либо 4, либо 5, либо 6 очков; вероятности этих четырех возможных результатов указаны в таблице (III) на стр. 79. Если спросить: «сколько очков выбивают наши два стрелка при одном (двойном) выстреле?», то на этот вопрос мы не сможем ответить, потому что разные выстрелы дают разные результаты. Но для оценки качества стрельбы нашей пары мы будем, конечно, интересоваться результатом не отдельного выстрела (этот результат может быть случайным), а *средним* результатом за целую серию выстрелов. Сколько же очков дает в среднем один выстрел нашей пары стрелков? Этот вопрос поставлен уже вполне разумно, и на него может быть дан ясный ответ.

Будем рассуждать так. Если наша пара стрелков производит сто двойных выстрелов, то, как показывает таблица (III),

примерно 4 из этих выстрелов дадут по 3 очка
» 26 » » » » » 4 »
» 46 » » » » » 5 очков
» 24 » » » » » 6 »

Таким образом, в среднем, каждая сотня двойных выстрелов даст нашей паре общее число очков, выражающееся суммой

$$3 \cdot 4 + 4 \cdot 26 + 5 \cdot 46 + 6 \cdot 24 = 490.$$

Деля это число на 100, мы получаем, что в среднем на один выстрел приходится 4,9 очка; это и дает ответ на поставленный нами вопрос. Заметим, что вместо того, чтобы делить на 100 готовую сумму (490) (как мы это только что делали), мы могли бы еще до сложения разделить на 100 каждое из слагаемых; тогда сумма прямо дает нам среднее число очков на один выстрел; проще всего произвести это деление, деля на 100 вторые множители всех слагаемых: ведь эти множители были получены умножением вероятностей, указанных в таблице (III), на 100, и для того чтобы разделить их на 100, достаточно просто вернуться к этим вероятностям. Для среднего числа очков, приходящегося на один выстрел, мы получаем, таким образом, выражение

$$3 \cdot 0{,}04 + 4 \cdot 0{,}26 + 5 \cdot 0{,}46 + 6 \cdot 0{,}24 = 4{,}9.$$

Сумма, стоящая в левой части этого равенства, как мы непосредственно видим, построена из данных таблицы (III) по очень простому правилу: каждое из указанных в верхней строчке этой таблицы возможных значений умножено на стоящую под ним в таблице его вероятность, и все такие произведения сложены между собой.

Проведем теперь это же рассуждение в общем виде. Допустим, что некоторая случайная величина задана таблицей

x_1	x_2	...	x_k
p_1	p_2	...	p_k

Вспомним: если вероятность значения x_1 величины x равна p_1, то это означает, что в серии из n операций это значение x_1 будет наблюдаться примерно n_1 раз, где $n_1/n = p_1$, откуда $n_1 = np_1$; аналогично, значение

x_2 при этом встретится примерно $n_2 = np_2$ раз, ... значение x_k встретится примерно $n_k = np_k$ раз.

Таким образом, серия из n операций будет в среднем содержать

$$
\begin{array}{lll}
n_1 = np_1 & \text{операций, где} & x = x_1, \\
n_2 = np_2 & \quad » \qquad » & x = x_2, \\
\dots & \dots & \dots \\
n_k = np_k & \text{операций, где} & x = x_k.
\end{array}
$$

Сумма значений величины x во всех n произведенных операциях будет поэтому примерно равна

$$x_1 n_1 + x_2 n_2 + \ldots + x_k n_k = n(x_1 p_1 + x_2 p_2 + \ldots + x_k p_k).$$

Поэтому среднее значение $\bar{x}$ случайной величины, соответствующее отдельной операции и получаемое из только что написанной суммы делением на число n операций в данной серии, будет равно

$$\bar{x} = x_1 p_1 + x_2 p_2 + \ldots + x_k p_k.$$

Мы приходим, таким образом, к следующему важному правилу: *для получения среднего значения случайной величины надо каждое из ее возможных значений помножить на соответствующую ему вероятность и сложить между собой все полученные произведения.*

Какую пользу может сослужить нам знание среднего значения случайной величины? Чтобы ответить на этот вопрос более убедительно, рассмотрим сначала несколько примеров.

Пример 1. Вернемся еще раз к нашим двум стрелкам; выбиваемые ими числа очков являются случайными величинами, законы распределения которых даются таблицами (I) для первого стрелка и (II) — для второго (стр. 77 и 78). Один внимательный взгляд на эти две таблицы уже показывает нам, что первый стреляет лучше второго; в самом деле, вероятность самого лучшего результата (3 очка) у него значительно больше, чем у второго стрелка, в то время как вероятности более плохих результатов, напротив, у второго стрелка больше, чем у первого. Однако такое сравнение не удовлетворяет нас — оно носит чисто качественный характер, при нем мы не видим еще той мерки, того числа, которое своей величиной прямо оценивало бы качество стрельбы того или другого стрелка, подобно тому как температура, например, прямо оценивает степень нагретости физического тела. Не имея такой оценивающей мерки, мы можем всегда стать перед случаем, когда непосредственное рассмотрение не даст нам никакого ответа или когда этот ответ может оказаться спорным. Так, если бы вместо таблиц (I) и (II) мы имели таблицы

1	2	3
0,4	0,1	0,5

(I′)

1	2	3
0,1	0,6	0,3

(II′)

((I′) — для первого стрелка, (II′) — для второго), то трудно было бы одним взглядом на эти таблицы решить, который из двух стрелков стреляет лучше; правда, лучший результат (3 очка) у первого более вероятен, чем у второго; но вместе с тем и самый худший результат (1 очко) у первого также более вероятен, чем у второго; напротив, результат 2 очка у второго более вероятен, чем у первого.

Составим теперь по указанному выше правилу среднее значение числа очков для каждого из наших двух стрелков:

1) для первого стрелка:

$$1 \cdot 0{,}4 + 2 \cdot 0{,}1 + 3 \cdot 0{,}5 = 2{,}1;$$

2) для второго стрелка:

$$1 \cdot 0{,}1 + 2 \cdot 0{,}6 + 3 \cdot 0{,}3 = 2{,}2.$$

Мы видим, что второй стрелок дает в среднем несколько большее число очков, чем первый; практически это значит, что при многократной стрельбе второй стрелок будет, вообще говоря, давать несколько лучший результат, чем первый. Теперь мы с уверенностью скажем, что второй стрелок стреляет лучше; среднее значение числа выбиваемых очков дало нам удобную мерку, с помощью которой мы можем легко и не оставляющим никаких сомнений способом сравнивать между собой искусство различных стрелков.

Пример 2. При сборке точного прибора для наиболее точной подгонки некоторой детали может потребоваться, в зависимости от удачи, 1, 2, 3, 4 или 5 проб. Таким образом, число x проб, необходимых для достижения удовлетворительной сборки, есть случайная величина с возможными значениями 1, 2, 3, 4, 5; пусть вероятности этих значений даются таблицей

1	2	3	4	5
0,07	0,16	0,55	0,21	0,01

Мы можем быть поставлены перед задачей снабдить данного сборщика таким числом деталей, какое необходимо для 20 приборов *). Чтобы иметь возможность ориентировочно оценить это число, мы не можем непосредственно использовать данную таблицу — она учит нас только тому, что в разных случаях бывает по-разному. Но если мы найдем среднее значение $\bar{x}$ числа x проб, необходимых для одного прибора, и умножим это среднее значение на 20, то мы и получим, очевидно, ориентировочное значение искомого числа. Мы находим:

$$\bar{x} = 1 \cdot 0{,}07 + 2 \cdot 0{,}16 + 3 \cdot 0{,}55 + 4 \cdot 0{,}21 + 5 \cdot 0{,}01 = 2{,}93;$$

$$20\bar{x} = 2{,}93 \cdot 20 = 58{,}6 \approx 59.$$

Для того чтобы сборщик имел небольшой запас на случай, если фактический расход деталей превзойдет ожидаемый, практически полезно будет дать ему 60—65 деталей.

В рассмотренных примерах мы имеем дело с таким положением вещей, когда для некоторой случайной величины практика требует известной ориентировочной оценки; одним только взглядом на таблицу мы такой оценки дать не можем, — таблица говорит нам

*) Мы будем предполагать при этом, что деталь, забракованная при сборке одного прибора, уже не используется при сборке других.

только, что наша случайная величина может принимать такие-то значения с такими-то вероятностями. Но вычисленное по этой таблице *среднее значение* случайной величины уже способно дать такую оценку, ибо это именно то значение, которое в среднем будет принимать наша величина при более или менее продолжительном ряде операций. Мы видим, что с практической стороны среднее значение особенно хорошо характеризует случайную величину, когда речь идет об операции массовой или многократно повторяемой.

Задача 1. Производится ряд испытаний с одной и той же вероятностью p появления некоторого события A, причем результаты отдельных испытаний между собой независимы. Найти среднее значение числа появлений события A в серии из n испытаний.

Число появлений события A в серии из n испытаний есть случайная величина с возможными значениями $0, 1, 2, \ldots, n$, причем вероятность значения k равна, как мы знаем,

$$P_n(k) = \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k}.$$

Поэтому искомое среднее значение равно

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

Эту сумму мы вычислили при доказательстве теоремы Бернулли (стр. 68) и видели что она равна np. В свое время мы убедились, что *наивероятнейшее* число появлений события A при n испытаниях в случае большого n близко к np; теперь мы видим, что *среднее* число появлений события A при любом n в точности равно np.

Таким образом, в данном случае наивероятнейшее значение случайной величины совпадает с ее средним значением; надо, однако, остерегаться думать, будто это совпадение имеет место для любых случайных величин; вообще говоря, наивероятнейшее значение случайной величины может очень далеко отстоять от ее среднего значения. Так, например, для случайной величины с законом распределения

0	5	10
0,7	0,1	0,2

наивероятнейшее значение есть 0, а среднее значение — 2,5.

Задача 2. Производятся независимые испытания, в каждом из которых с вероятностью 0,8 может произойти некоторое событие A. Испытания производятся до первого появления события A; общее число испытаний не превосходит четырех. Определить среднее число произведенных испытаний.

Число испытаний, которое придется произвести, по условиям задачи может равняться 1, 2, 3 или 4. Мы должны вычислить вероятности каждого из этих четырех значений. Для того чтобы потребовалось произвести только одно испытание, надо, чтобы уже при первом испытании появилось событие A. Вероятность этого равна

$$p_1 = 0{,}8.$$

Для того чтобы потребовалось произвести два испытания, надо, чтобы при первом испытании событие A не появилось, а при втором — произошло. Вероятность этого по теореме умножения вероятностей для независимых событий равна

$$p_2 = (1 - 0{,}8) \cdot 0{,}8 = 0{,}16.$$

Для того чтобы потребовалось три испытания, надо, чтобы в первых двух событие A не появилось, а при третьем оно произошло.

Поэтому

$$p_3 = (1 - 0{,}8)^2 \cdot 0{,}8 = 0{,}032.$$

Наконец, потребность в четырех испытаниях возникнет при условии, что три первых испытания не приведут к появлению события A (независимо от того, что даст четвертое испытание), поэтому

$$p_4 = (1 - 0{,}8)^3 = 0{,}008.$$

Таким образом, число производимых испытаний, как случайная величина, определяется законом распределения

1	2	3	4
0,8	0,16	0,032	0,008

Среднее значение этого числа равно поэтому

$$1 \cdot 0{,}8 + 2 \cdot 0{,}16 + 3 \cdot 0{,}032 + 4 \cdot 0{,}008 = 1{,}248.$$

Если, например, предстоит сделать 100 подобных наблюдений, то можно рассчитывать, что при этом придется произвести примерно $1{,}248 \cdot 100 \approx 125$ испытаний.

В практике с подобной постановкой задачи приходится встречаться часто. Для примера, мы испытываем пряжу на крепость и относим ее к высшему сорту, если она не рвется ни разу при нагрузке P, когда испытываются образцы стандартной длины из одного и того же мотка (или партии). Испытывается каждый раз не более четырех образцов.

Задача 3. Некоторая площадка имеет форму квадрата, сторона которого по данным аэрофотометрических измерений равна 350 *м*. Качество аэрофотосъемки определяется тем, что ошибка в

0 *м*	имеет	вероятность	0,42
± 10 *м*	»	»	0,16 *)
± 20 *м*	»	»	0,08
± 30 *м*	»	»	0,05

Найти среднее значение площади площадки.

В зависимости от случайностей аэрофотометрического измерения сторона площадки есть случайная

*) Это понимается так, что ошибка +10 *м* и ошибка —10 *м* имеют вероятности по 0,16 каждая; так же следует понимать и далее указанные вероятности.

величина, закон распределения которой дается таблицей

320	330	340	350	360	370	380
0,05	0,08	0,16	0,42	0,16	0,08	0,05

(I)

Отсюда мы сразу могли бы найти среднее значение этой величины; в данном случае нам нет даже надобности применять для этого наше вычислительное правило; в самом деле, так как одинаковые ошибки в ту и другую сторону одинаково вероятны, то уже по симметрии ясно, что среднее значение стороны квадрата равно наблюденному значению, т. е. 350 *м*; более подробно: выражение среднего значения будет содержать члены

$$(340+360)\cdot 0,16=[(350+10)+(350-10)]\cdot 0,16= \\ =2\cdot 350\cdot 0,16,$$

$$(330+370)\cdot 0,08=2\cdot 350\cdot 0,08,$$

$$(320+380)\cdot 0,05=2\cdot 350\cdot 0,05;$$

оно равно поэтому

$$350\cdot(0,42+2\cdot 0,16+2\cdot 0,08+2\cdot 0,05)=350.$$

Можно было бы думать, что из тех же соображений симметрии среднее значение площади квадрата должно равняться $350^2=122\,500$ *м*^{2}; это было бы так, если бы среднее значение квадрата случайной величины равнялось квадрату ее среднего значения, однако это не так; в нашем примере площадь квадрата может иметь значения

$$320^2,\ 330^2,\ 340^2,\ 350^2,\ 360^2,\ 370^2,\ 380^2;$$

какое из этих значений имеет место в действительности — это зависит от того, какой из семи представленных в таблице (I) случаев имеется налицо, так что вероятности этих семи значений те же, что и вероятности таблицы (I); короче говоря, закон распределения площади квадрата дается таблицей

320^2	330^2	340^2	350^2	360^2	370^2	380^2
0,05	0,08	0,16	0,42	0,16	0,08	0,05

и, следовательно, среднее значение ее равно

$$320^2 \cdot 0{,}05 + 330^2 \cdot 0{,}08 + 340^2 \cdot 0{,}16 + 350^2 \cdot 0{,}42 + \\ + 360^2 \cdot 0{,}16 + 370^2 \cdot 0{,}08 + 380^2 \cdot 0{,}05.$$

И здесь полезно для сокращения вычислений воспользоваться имеющейся симметрией; стоит посмотреть, как это делается, ибо подобные возможности упрощения возникают довольно часто. Мы можем переписать написанное выражение в виде

$$\begin{aligned} &350^2 \cdot 0{,}42 + (340^2 + 360^2) \cdot 0{,}16 + (330^2 + 370^2) \cdot 0{,}08 + \\ &\qquad + (320^2 + 380^2) \cdot 0{,}05 = \\ &= 350^2 \cdot 0{,}42 + [(350-10)^2 + (350+10)^2] \cdot 0{,}16 + \\ &\qquad + [(350-20)^2 + (350+20)^2] \cdot 0{,}08 + \\ &\qquad + [(350-30)^2 + (350+30)^2] \cdot 0{,}05 = \\ &= 350^2 [0{,}42 + 2 \cdot 0{,}16 + 2 \cdot 0{,}08 + 2 \cdot 0{,}05] + \\ &\qquad + 2 \cdot 10^2 \cdot 0{,}16 + 2 \cdot 20^2 \cdot 0{,}08 + 2 \cdot 30^2 \cdot 0{,}05 = \\ &= 350^2 + 2 \cdot (16 + 32 + 45) = 122\,686. \end{aligned}$$

При этом способе вычисления все расчеты могут быть проведены в уме.

Мы видим, что среднее значение площади квадрата оказалось несколько больше (практически, правда, различие в этом случае неощутимо), чем квадрат среднего значения стороны (т. е. $350^2 = 122\,500$); легко доказать, что в основе этого лежит общее правило: среднее значение квадрата любой случайной величины всегда больше, чем квадрат ее среднего значения. В самом деле, пусть мы имеем случайную величину x с совершенно произвольным законом распределения

x_1	x_2	...	x_k
p_1	p_2	...	p_k

тогда величина x^2 будет иметь закон распределения

x_1^2	x_2^2	...	x_k^2
p_1	p_2	...	p_k

средние значения этих двух величин соответственно равны

$$\bar{x} = x_1p_1 + x_2p_2 + \ldots + x_kp_k$$

и

$$\bar{x^2} = x_1^2p_1 + x_2^2p_2 + \ldots + x_k^2p_k.$$

Мы имеем:

$$\bar{x^2} - (\bar{x})^2 = \bar{x^2} - 2(\bar{x})^2 + (\bar{x})^2;$$

так как $p_1 + p_2 + \ldots + p_k = 1$, то три члена правой части могут быть соответственно представлены в виде

$$\bar{x^2} = \sum_{i=1}^{k} x_i^2 p_i,$$

$$2(\bar{x})^2 = 2(\bar{x})(\bar{x}) = 2\bar{x}\sum_{i=1}^{k} x_ip_i = \sum_{i=1}^{k} 2\bar{x}x_ip_i,$$

$$(\bar{x})^2 = (\bar{x})^2\sum_{i=1}^{k} p_i = \sum_{i=1}^{k} (\bar{x})^2 p_i;$$

поэтому

$$\bar{x^2} - (\bar{x})^2 = \sum_{i=1}^{k} \{x_i^2 - 2\bar{x}x_i + (\bar{x})^2\} p_i = \sum_{i=1}^{k} (x_i - \bar{x})^2 p_i;$$

так как все члены суммы, стоящей в правой части, неотрицательны, то

$$\bar{x^2} - (\bar{x})^2 > 0,$$

что и требовалось доказать.

ГЛАВА ДЕВЯТАЯ

СРЕДНИЕ ЗНАЧЕНИЯ СУММЫ И ПРОИЗВЕДЕНИЯ

§ 21. Теорема о среднем значении суммы

Очень часто встречается необходимость вычислить среднее значение суммы двух (а иногда и большего числа) случайных величин, средние значения которых известны. Пусть, например, два предприятия изготовляют одинаковую продукцию, причем известно, что в среднем первое предприятие ежедневно производит 120 изделий, а второе — 180; можем ли мы при помощи этих данных установить среднее значение числа изделий, которое следует ожидать ежедневно от обоих предприятий вместе? Или этих данных недостаточно, и мы, кроме средних значений, должны знать еще что-либо о двух рассматриваемых случайных величинах (например, знать полностью их законы распределения)?

Весьма важно, что для вычисления среднего значения суммы во всех случаях достаточно знать средние значения слагаемых и что среднее значение суммы выражается во всех случаях через средние значения слагаемых наиболее простым способом, какой только можно себе представить: *среднее значение суммы всегда равно сумме средних значений слагаемых.*

Таким образом, если x и y — две совершенно произвольные случайные величины, то

$$\overline{x+y}=\bar{x}+\bar{y}.$$

В вышеприведенном примере x — число изделий одного только первого предприятия, y — число изделий второго предприятия: $\bar{x} = 120$, $\bar{y} = 180$, и значит,

$$\overline{x+y} = \bar{x} + \bar{y} = 300.$$

Чтобы доказать в общем виде утверждаемое правило, допустим, что величины x и y соответственно подчиняются законам распределения

x_1	x_2	...	x_k
p_1	p_2	...	p_k

(I)

y_1	y_2	...	y_l
q_1	q_2	...	q_l

(II)

Тогда возможными значениями величины $x + y$ будут всевозможные суммы вида $x_i + y_j$, где $1 \leqslant i \leqslant k$ и $1 \leqslant j \leqslant l$. Вероятность значения $x_i + y_j$, которую будем обозначать через p_{ij}, нам неизвестна; это есть вероятность двойного события $x = x_i$, $y = y_j$, т. е. вероятность того, что величина x получит значение x_i, а величина y — значение y_j. Если бы мы могли считать эти два события взаимно независимыми, то по правилу умножения имели бы, конечно,

$$p_{ij} = p_i q_j, \tag{1}$$

но мы отнюдь не станем предполагать независимости этих событий.

Итак, равенство (1), вообще говоря, не будет иметь места, и нужно считаться с тем, что знание таблиц (I) и (II) ничего не позволяет заключить о величинах p_{ij}.

По общему правилу, среднее значение величины $x + y$ равно сумме произведений всех возможных значений этой величины на соответствующие вероятности, т. е.

$$\overline{x+y} = \\ = \sum_{i=1}^{k} \sum_{j=1}^{l} (x_i + y_j) p_{ij} = \sum_{i=1}^{k} x_i \left(\sum_{j=1}^{l} p_{ij} \right) + \sum_{j=1}^{l} y_j \left(\sum_{i=1}^{k} p_{ij} \right). \tag{2}$$

Рассмотрим теперь внимательнее сумму $\sum_{j=1} p_{ij}$; это есть сумма вероятностей всевозможных событий вида $(x = x_i,\ y = y_j)$, где число i — одно и то же во всех членах суммы, а число j от члена к члену меняется, пробегая весь ряд возможных для него значений (от 1 до l включительно); так как события $y = y_j$ при различных j друг с другом, очевидно, несовместимы, то по правилу сложения сумма $\sum_{j=1}^{l} p_{ij}$ есть вероятность наступления *какого-нибудь одного из l событий* $(x = x_i,\ y = y_j)$ $(j = 1, 2, \ldots, l)$.

Но сказать «наступило какое-нибудь одно из событий $x = x_i,\ y = y_j$ $(1 \leqslant j \leqslant l)$» — все равно, что просто сказать «наступило событие $x = x_i$»; в самом деле: 1) если наступило одно из событий $(x = x_i,\ y = y_j)$ (j — все равно какое), то, очевидно, наступило и событие $x = x_i$; 2) обратно, если наступило событие $x = x_i$, то, так как y обязательно принимает одно из своих возможных значений $y_1, y_2, \ldots, y_l$, должно наступить и какое-нибудь из событий $(x = x_i,\ y = y_j)$ $(1 \leqslant j \leqslant l)$. Таким образом, $\sum_{j=1} p_{ij}$, будучи вероятностью наступления какого-нибудь одного из событий $(x = x_i,\ y = y_j)$ $(1 \leqslant j \leqslant l)$, равна просто вероятности события $x = x_i$, т. е.

$$\sum_{j=1}^{l} p_{ij} = p_i.$$

Совершенно аналогичным образом мы убеждаемся, конечно, что

$$\sum_{i=1}^{k} p_{ij} = q_j;$$

вставляя же эти выражения в равенство (2), мы находим:

$$\overline{x+y} = \sum_{i=1}^{k} x_i p_i + \sum_{j=1}^{l} y_j q_j = \bar{x} + \bar{y},$$

что и требовалось доказать.

Теорема, доказанная нами для случая двух слагаемых, автоматически распространяется на случай трех и более слагаемых; в самом деле, в силу того, что мы только что доказали, что можем написать

$$\overline{x+y+z}=\overline{x+y}+\bar{z}=\bar{x}+\bar{y}+\bar{z}$$

и т. д.

Пример. На некотором предприятии установлено n станков и с каждого станка отобрано по одному изделию. Определить среднее число бракованных изделий, если известно, что вероятность изготовления бракованного изделия первым станком равна p_1, вторым станком — p_2, ..., n-м станком — p_n.

Число бракованных изделий при анализе одного изделия есть случайная величина, способная принимать только два значения: 1, если это изделие бракованное, и 0, если оно годное. Вероятности этих значений для первого станка равны соответственно p_1 и $1-p_1$, вследствие чего среднее число бракованных изделий из числа взятых с первого станка равно

$$1\cdot p_1+0\cdot(1-p_1)=p_1.$$

Для второго станка среднее число бракованных изделий среди взятых равно p_2 и т. д. Общее число бракованных изделий есть сумма бракованных изделий, попавшихся среди изделий, изготовленных на первом, втором и других станках. Поэтому, в силу только что установленного нами правила сложения средних значений, среднее число бракованных изделий среди выбранных нами равно

$$p_1+p_2+\ldots+p_n,$$

что и решает поставленную задачу.

В частности, если вероятность изготовления бракованного изделия одна и та же для всех станков ($p_1=p_2=\ldots=p_n=p$), то среднее значение общего числа бракованных изделий равно np. Этот результат мы уже получили на стр. 68 [формула (5)]. Интересно сравнить громоздкие расчеты, которые понадобились там для этой цели, с тем простым, не требующим никаких вычислений рассуждением, которое здесь привело нас к тому же результату. Однако мы выгадали не только в простоте, но и в общности. При нашем прежнем выводе мы предполагали результаты изготовления отдельных изделий взаимно **независимыми** событиями, и действительно, наш прежний способ вывода годился только при этом предположении; теперь же мы можем обойтись без этого предположения, так как правило сложения средних значений, на которое мы опираемся в нашем новом выводе, имеет место для любых случайных величин, без всякого ограничения. Таким образом, какова бы ни была взаимная зависимость между отдельными станками и изготовляемыми ими изделиями, если только вероятность изготовления брака p для всех станков одна и та же, среднее число бракованных изделий среди выбранных n всегда равно np.

§ 22. Теорема о среднем значении произведения

Тот же вопрос, который мы разрешили для сумм случайных величин, часто приходится рассматривать и для их произведений. Пусть случайные величины x и y снова подчиняются соответственно законам распределения, указываемым таблицами (I) и (II). Тогда произведение xy есть случайная величина, возможными значениями которой служат произведения вида x_iy_j $(1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant l)$; вероятность значения x_iy_j равна p_{ij}. Задача состоит в отыскании такого правила, которое позволило бы во всех случаях выражать среднее значение $\overline{xy}$ величины xy через средние значения сомножителей. Однако в общем случае решение этой задачи оказывается невозможным. Знанием средних значений $\bar{x}$ и $\bar{y}$ величина $\overline{xy}$, вообще говоря, не определяется однозначно (т. е. при одних и тех же $\bar{x}$ и $\bar{y}$ возможны различные значения величины $\overline{xy}$); вследствие этого никакой общей формулы, выражающей $\overline{xy}$ через $\bar{x}$ и $\bar{y}$, существовать не может.

Но есть один важный случай, когда такое выражение возможно, и тогда получаемая связь носит чрезвычайно простой характер.

Условимся называть случайные величины x и y *взаимно независимыми*, если события $x = x_i$ и $y = y_j$

при любых i и j взаимно независимы, т. е. если условие, что одна из наших двух случайных величин приняла то или другое определенное значение, не влияет на закон распределения второй случайной величины. Если величины x и y взаимно независимы в только что определенном смысле, то

$$p_{ij} = p_i q_j \qquad (i = 1,\ 2,\ \ldots,\ k;\ j = 1,\ 2,\ \ldots,\ l)$$

по правилу умножения для независимых событий; поэтому

$$\overline{xy} = \sum_{i=1}^{k} \sum_{j=1}^{l} x_i y_j p_{ij} = \sum_{i=1}^{k} \sum_{j=1}^{l} x_i y_j p_i q_j = \sum_{i=1}^{k} x_i p_i \sum_{j=1}^{l} y_j q_j = \bar{x} \cdot \bar{y};$$

для взаимно независимых случайных величин среднее значение произведения равно произведению средних значений сомножителей.

Как и в случае сложения, это правило, выведенное нами для произведения двух случайных величин, автоматически распространяется на произведение любого числа сомножителей; при этом необходимо только, чтобы эти сомножители были взаимно независимыми, т. е. чтобы задание каких-либо определенных значений для части этих величин не влияло на законы распределения остальных величин.

Пример 1. Допустим, что требуется измерить площадь прямоугольной формы посредством аэрофотосъемки и что измерение сторон этого прямоугольника дало 72 *м* и 50 *м*. Закон распределения ошибок измерения неизвестен, но известно, что ошибки одинаковой величины в ту и другую сторону одинаково вероятны; тогда ясно из соображений симметрии (и легко может быть доказано, см. задачу 3 на стр. 87), что средние значения сторон прямоугольника совпадают с полученными результатами измерения. Если эти два результата измерения можно считать взаимно независимыми случайными величинами, то среднее значение площади прямоугольника по только что выведенному нами правилу умножения будет равно произведению средних значений его сторон, т. е. $72 \cdot 50 = 3600$ $м^2$. Но иногда могут быть основания предполагать измерения сторон взаимно зависимыми. Это будет, например, в том случае, если оба измерения производятся одними и теми же недостаточно выверенными приборами. Если измерение длины дает результат, значительно превосходящий истинную длину, то мы, естественно, имеем основание предполагать, что измерительные приборы вообще склонны давать слишком большие величины, вследствие чего повышается вероятность преувеличенных значений и при измерении ширины, так что эти две величины нельзя считать взаимно независимыми. В таких случаях среднее значение площади не может быть принято равным произведению средних значений сторон прямоугольника, и для определения его требуется дополнительная информация.

Пример 2. По проводнику, сопротивление которого зависит от случайных обстоятельств, течет электрический ток, сила которого также зависит от случая. Известно, что среднее значение сопротивления проводника равно 25 омам, а средняя сила тока равна 6 амперам. Требуется подсчитать среднее значение электродвижущей силы E текущего по проводнику тока.

Согласно закону Ома

$$E = RI,$$

где R — сопротивление проводника, I — сила тока.

Так как по нашему допущению

$$\overline{R} = 25, \quad \overline{I} = 6,$$

то, предположив величины R и I взаимно независимыми, находим, что

$$\overline{E} = \overline{R}\overline{I} = 25 \cdot 6 = 150 \text{ вольт.}$$

ГЛАВА ДЕСЯТАЯ

РАССЕЯНИЕ И СРЕДНИЕ УКЛОНЕНИЯ

§ 23. Недостаточность среднего значения для характеристики случайной величины

Мы уже неоднократно видели, что среднее значение случайной величины дает нам первое примерное, ориентировочное представление о ней, и есть много случаев, когда для практических целей, стоящих перед нами, этого представления бывает достаточно. Так, для сравнения искусства двух стрелков в соревновании нам достаточно было знать средние значения чисел выбиваемых ими очков; для сравнения эффективности двух различных систем подсчета числа космических частиц вполне достаточно знать средние значения числа потерь частиц, которые допускает каждая из систем измерений, и т. д. Во всех этих случаях мы получаем существенную выгоду, описывая нашу случайную величину одним числом — ее средним значением, — вместо того чтобы задавать ее сложным законом распределения. Дело обстоит так, как будто бы перед нами была не случайная, а достоверно известная величина с совершенно определенным значением.

Однако гораздо чаще встречается другое положение вещей, когда наиболее важные для практических целей черты случайной величины ни в какой мере не определяются ее средним значением, а требуют более детального знакомства с ее законом распределения. Типичный случай этого рода мы имеем при исследовании распределения ошибок измерений. Пусть x — величина ошибки, т. е. отклонение полученного значения измеряемой величины от ее истинного значения. При отсутствии систематических ошибок среднее значение ошибки измерения, которое мы обозначим через $\bar{x}$, равно нулю. Допустим, что измерения происходят именно при таком условии. Спрашивается, как будут распределяться ошибки? Как часто будут встречаться ошибки той или иной величины? На все эти вопросы мы, зная только значение $\bar{x} = 0$, не можем получить никакого ответа. Мы знаем только, что возможны ошибки положительные и отрицательные и что их шансы примерно одинаковы, ибо среднее значение величины ошибки равно нулю. Но мы не знаем самого главного: будут ли результаты измерений в своем большинстве ложиться близко к истинному значению измеряемой величины, так что можно будет рассчитывать с большой надежностью на каждый результат измерения, или же они в своей основной массе будут раскиданы на больших расстояниях от истинного размера по обе стороны. Обе возможности вполне допустимы.

Два наблюдателя могут, производя измерения с одним и тем же средним значением ошибки $\bar{x}$, давать измерения различной степени точности. Может случиться, что один из них дает систематически большее «рассеяние» результатов измерений, чем другой. Это означает, что у этого наблюдателя ошибки в среднем могут принимать бо́льшие значения, а значит, измерения будут дальше отстоять от измеряемой величины, чем у другого. И это возможно, хотя для обоих наблюдателей средняя величина ошибки измерения одна и та же.

Рассмотрим другой пример. Представим себе, что испытываются на урожайность два сорта пшеницы. В зависимости от случайных обстоятельств (количество осадков, распределение удобрений, солнечная радиация и пр.) урожай с квадратного метра подвержен значительным колебаниям и представляет собой случайную величину. Предположим, что при одинаковых условиях средний урожай для каждого сорта один и тот же — 240 граммов с квадратного метра. Можно ли судить о качестве испытываемого сорта только по значению среднего урожая? Очевидно, что нет, так как наибольший хозяйственный интерес представляет тот сорт, урожайность которого меньше подвержена случайным влияниям метеорологических и других факторов, иными словами, для которого «рассеяние» урожайности меньше. Мы видим, таким образом, что при испытании того или иного сорта пшеницы на урожайность не меньшее значение, чем средняя урожайность, имеет размах возможных ее колебаний.

§ 24. Различные способы измерения рассеяния случайной величины

Приведенные примеры, так же как и ряд других, им аналогичных, убедительно показывают, что во многих случаях для описания практически наиболее интересных черт случайных величин задания их средних значений бывает совершенно недостаточно. Эти практически интересные черты при таком задании остаются неизвестными, и для освещения их мы должны либо иметь перед собой всю таблицу распределения такой величины, что практически почти всегда сложно и неудобно, либо постараться, кроме среднего значения такой величины, ввести для ее описания еще одно-два числа подобного же рода, так чтобы в своей совокупности эта небольшая группа чисел давала практически достаточную характеристику тех черт изучаемой величины, которые представляются нам наиболее существенными. Посмотрим, как может быть осуществлена эта последняя возможность.

Как показывают рассмотренные нами примеры, во многих случаях практически наиболее важным оказывается вопрос о том, насколько велики, вообще говоря, отклонения значений, фактически принимаемых данной случайной величиной, от ее среднего значения, т. е. как сильно эти значения раскиданы, рассеяны; будут ли они по большей части тесно сгруппированы вокруг среднего значения (а значит, и между собой), или, напротив, большинство из них будет отличаться от среднего значения очень резко (в этом случае некоторые из них по необходимости будут значительно отличаться и друг от друга).

Следующая грубая схема позволяет ясно представить себе картину этого различия. Рассмотрим две случайные величины со следующими распределениями вероятностей:

−0,01	+0,01
0,5	0,5

(I)

−100	+100
0,5	0,5

(II)

Обе случайные величины, таблицы которых мы выписали, имеют своим средним значением нуль, но, в то время как первая из них всегда принимает значения, очень близкие к нулю (и между собой), вторая, наоборот, способна принимать лишь значения, резко отличающиеся от нуля (и друг от друга). Для первой величины знание ее среднего значения дает нам вместе с тем и ориентировочные сведения о ее фактических возможных значениях; для второй же — среднее значение удалено от фактических возможных значений весьма значительно и не дает о них никакого представления. Мы говорим, что во втором случае возможные значения *рассеяны* гораздо больше, чем в первом.

Таким образом, наша задача — найти число, которое разумным образом могло бы давать нам *меру* рассеяния случайной величины, которое хотя бы ориентировочно указывало нам, сколь больших отклонений этой величины от ее среднего значения нам следует ожидать. Отклонение $x - \bar{x}$ случайной величины x от ее среднего значения $\bar{x}$ само есть, очевидно, случайная величина; случайной величиной будет и абсолютное значение $|x - \bar{x}|$ этого отклонения, характеризующее его размер вне зависимости от знака. Желательно иметь число, которое могло бы ориентировочно охарактеризовать это случайное отклонение $|x - \bar{x}|$, сказать нам о том, сколь большим, примерно, может оказаться это отклонение. Для решения этого вопроса существует много различных способов, из которых наиболее употребительны на практике следующие три:

1. С р е д н е е у к л о н е н и е. За ориентировочное значение случайной величины $|x - \bar{x}|$ естественнее

всего принять ее среднее значение $\overline{|x-\bar{x}|}$. Это среднее значение абсолютной величины уклонения называют *средним уклонением* величины x. Если случайная величина x задается таблицей

x_1	x_2	...	x_k
p_1	p_2	...	p_k

то таблица случайной величины $|x-\bar{x}|$ имеет вид

$\lvert x_1-\bar{x}\rvert$	$\lvert x_2-\bar{x}\rvert$	...	$\lvert x_k-\bar{x}\rvert$
p_1	p_2	...	p_k

где $\bar{x}=\sum_{i=1}^{k} x_i p_i$; для среднего уклонения M_x величины x получаем формулу

$$M_x=\overline{|x-\bar{x}|}=\sum_{i=1}^{k} |x_i-\bar{x}|\, p_i,$$

где, разумеется, снова $\bar{x}=\sum_{i=1}^{k} x_i p_i$. Для величин, заданных таблицами (I) и (II), $\bar{x}=0$, и мы соответственно имеем:

$$M_{x_{\mathrm{I}}}=0{,}01 \quad \text{и} \quad M_{x_{\mathrm{II}}}=100.$$

Впрочем, оба примера тривиальны, так как в обоих случаях абсолютная величина уклонения оказывается способной принимать только одно значение, утрачивая, таким образом, характер случайной величины.

Вычислим еще средние уклонения для случайных величин, определяемых таблицами (I′) и (II′) на стр. 83. Мы видели там, что средние значения этих величин соответственно равны 2,1 и 2,2, т. е. очень близки друг к другу. Среднее уклонение для первой величины равно

$$|1-2{,}1|\cdot 0{,}4+|2-2{,}1|\cdot 0{,}1+|3-2{,}1|\cdot 0{,}5=0{,}9,$$

а для второй —

$$|1-2{,}2|\cdot 0{,}1+|2-2{,}2|\cdot 0{,}6+|3-2{,}2|\cdot 0{,}3=0{,}48.$$

Для второй величины среднее уклонение почти вдвое меньше, чем для первой; практически это означает, очевидно, что хотя в среднем оба стрелка выбивают примерно одинаковое число очков и в этом смысле могут быть признаны одинаково искусными, но в то же время у второго стрелка стрельба носит значительно *более ровный* характер, его результаты значительно менее рассеяны, чем у первого стрелка, который при том же самом среднем числе очков стреляет неровно, часто давая результаты как много лучше, так и много хуже средних.

2. Среднее квадратическое уклонение. Измерять ориентировочную величину уклонения с помощью среднего уклонения — очень естественно, но вместе с тем и очень неудобно практически, так как вычисления и оценки с абсолютными величинами часто бывают сложны, а иногда и совсем недоступны. Поэтому на практике для величины уклонения обычно предпочитают вводить другую меру.

Так же, как и уклонение $x-\bar{x}$ случайной величины x от ее среднего значения $\bar{x}$, квадрат $(x-\bar{x})^2$ этого уклонения есть случайная величина, таблица которой в наших старых обозначениях имеет вид

$(x_1-\bar{x})^2$	$(x_2-\bar{x})^2$	...	$(x_k-\bar{x})^2$
p_1	p_2	...	p_k

среднее значение этого квадрата равно поэтому

$$\sum_{i=1}^{k}(x_i-\bar{x})^2 p_i.$$

Эта величина дает нам представление о том, чему примерно равен *квадрат* уклонения $x-\bar{x}$; поэтому, извлекая из этой величины квадратный корень

$$Q_x=\sqrt{\sum_{i=1}^{k}(x_i-\bar{x})^2 p_i},$$

мы получаем величину, способную охарактеризовать нам примерный размер самого уклонения. Величина Q_x называется *средним квадратическим уклонением* случайной величины x, а квадрат ее, т. е. величина Q_x^2, — ее *дисперсией*. Разумеется, эта новая мера величины уклонения носит характер несколько более искусственный, чем среднее уклонение, введенное нами выше: здесь мы идем обходным путем, находя сначала ориентировочное значение для *квадрата* уклонения и лишь потом, с помощью извлечения квадратного корня, возвращаясь к самому уклонению. Но зато, как мы увидим в следующем параграфе, использование средних квадратичных уклонений Q_x значительно упрощает вычисления. Именно это и заставляет статистиков на практике преимущественно пользоваться средними квадратическими уклонениями.

Пример. Для случайных величин, определяемых таблицами (I′) и (II′) на стр. 83, мы имеем соответственно:

$$Q_{x\mathrm{I}'}^2 = (1-2{,}1)^2\cdot 0{,}4 + (2-2{,}1)^2\cdot 0{,}1 + \\ + (3-2{,}1)^2\cdot 0{,}5 = 0{,}89$$

и

$$Q_{x\mathrm{II}'}^2 = (1-2{,}2)^2\cdot 0{,}1 + (2-2{,}2)^2\cdot 0{,}6 + \\ + (3-2{,}2)^2\cdot 0{,}3 = 0{,}36,$$

и, следовательно,

$$Q_{x\mathrm{I}'} = \sqrt{0{,}89} \approx 0{,}94 \quad \text{и} \quad Q_{x\mathrm{II}'} = 0{,}6.$$

Выше мы имели для тех же величин средние уклонения:

$$M_{x\mathrm{I}'} = 0{,}9 \quad \text{и} \quad M_{x\mathrm{II}'} = 0{,}48.$$

Мы видим, что среднее квадратическое уклонение, так же как и среднее уклонение, для первой величины значительно больше, чем для второй; будем ли мы измерять рассеяние средним или средним квадратическим уклонением, в обоих случаях мы приходим к тому же выводу, что первая из наших двух величин рассеяна в большей степени, чем вторая.

В обоих случаях у нас среднее квадратическое уклонение оказалось больше среднего уклонения; легко сообразить, что так оно должно быть для любой случайной величины. В самом деле, дисперсия Q_x^2, как среднее значение квадрата величины $|x-\bar{x}|$, по правилу, доказанному на стр. 89, не может быть меньше, чем квадрат среднего значения M_x величины $|x-\bar{x}|$, а из $Q_x^2 \geqslant M_x^2$ вытекает $Q_x \geqslant M_x$.

3. Срединное (вероятное) уклонение. Часто, особенно в военном деле, употребляется иной способ для определения размеров рассеивания; мы изложим его в терминах артиллерийского примера.

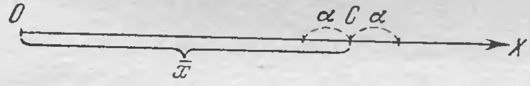

Рис. 9.

Допустим, что ведется артиллерийская стрельба из точки O в направлении OX (рис. 9); расстояние x места падения снаряда от места вылета есть случайная величина, среднее значение которой указывает нам положение «центра попаданий» C $(OC=\bar{x})$, вокруг которого более или менее рассеиваются точки падения отдельных снарядов.

Уклонение $x-\bar{x}$ изучаемой нами случайной величины (дальности полета снаряда) от ее среднего значения есть вместе с тем уклонение точки падения снаряда от центра попаданий C; всякая оценка величины $|x-\bar{x}|$ поэтому оценивает вместе с тем степень рассеяния, разброса снарядов вокруг этого центра C и служит, таким образом, важнейшим показателем качества стрельбы.

Если мы отложим от точки C влево и вправо по очень малому отрезку длины α, то лишь немногие снаряды будут ложиться внутри полученного таким образом отрезка длины 2α с серединой в точке C, иначе говоря, вероятность того, что $|x-\bar{x}|<\alpha$, при малых α будет еще очень мала. Но будем теперь расширять построенный нами отрезок, увеличивая число α (которое ведь было выбрано произвольно). Чем больше будет построенный отрезок, тем большая доля снарядов будет падать внутри него и тем больше будет, значит, для отдельного снаряда вероятность попасть внутрь этого отрезка; когда α очень велико, то практически все снаряды будут падать внутри этого отрезка; таким образом, с постепенным возрастанием числа α вероятность неравенства

$$|x-\bar{x}|<\alpha$$

растет от нуля до единицы; сначала, при малых α, вероятнее, что

$$|x-\bar{x}|>\alpha,$$

т. е. что снаряд упадет вне отрезка; а потом, при больших α, вероятнее, что будет $|x-\bar{x}|<\alpha$, т. е. что снаряд упадет внутри отрезка. Поэтому где-то при переходе от малых значений числа α к более крупным должно быть такое значение α_0 этого числа α, что снаряд имеет одинаковую вероятность упасть как внутри, так и вне отрезка длины $2\alpha_0$, построенного вокруг точки C. Другими словами, неравенства

$$|x-\bar{x}|<\alpha_0$$

и

$$|x-\bar{x}|>\alpha_0$$

одинаково вероятны, и, значит, вероятность каждого из них равна $1/2$ (если мы условимся пренебрегать ничтожно малой вероятностью точного *равенства* $|x-\bar{x}|=\alpha$). При $\alpha<\alpha_0$ более вероятно второе, при $\alpha>\alpha_0$ — первое из написанных неравенств. Таким образом, существует единственное определенное число α_0 такое, что абсолютная величина уклонения с одинаковой вероятностью может оказаться как больше, так и меньше чем α_0.

Как велико α_0 — это зависит от качеств стреляющего орудия; легко видеть, что величина α_0 может служить мерой рассеяния снарядов, подобно среднему или среднему квадратическому уклонению. В самом деле, если, например, α_0 очень мало, то это означает, что уже на очень малый окружающий точку C отрезок ложится половина всех выпускаемых орудием снарядов, что свидетельствует о сравнительно незначительном рассеянии. Напротив, если α_0 велико, то, окружив точку C даже большим отрезком, мы все же должны рассчитывать, что половина снарядов будет ложиться за пределами этого отрезка; а это, очевидно, показывает, что снаряды сильно рассеиваются вокруг центра.

Число α_0 называют обычно *срединным* или *вероятным уклонением* величины x; таким образом, срединным или вероятным уклонением случайной величины x мы называем такое число, что уклонение $x-\bar{x}$ с одинаковой вероятностью может оказаться по абсолютному значению как больше, так и меньше этого числа. Хотя срединное уклонение величины x, которое мы будем обозначать через E_x, для математических расчетов не более удобно, чем среднее уклонение M_x, и значительно менее удобно, чем среднее квадратическое уклонение Q_x, тем не менее в артиллерии принято для оценки всех уклонений пользоваться именно величиной E_x. В дальнейшем мы узнаем, почему это на практике не ведет обычно ни к каким затруднениям.

§ 25. Теоремы о среднем квадратическом уклонении

Убедимся теперь, что средние квадратические уклонения действительно обладают особыми свойствами, заставляющими предпочитать их всяким другим характеристикам величины уклонения — средним, срединным (вероятным) уклонениям и т. п. Как мы убедимся немного позднее, для приложений основное значение имеет следующая задача.

Пусть мы имеем случайные величины $x_1, x_2, \ldots, x_n$ со средними квадратическими уклонениями $q_1, q_2, \ldots, q_n$. Положим $x_1+x_2+\ldots+x_n=X$ и спросим себя, как найти среднее квадратическое уклонение Q величины X, если даны $q_1, q_2, \ldots, q_n$ и если мы предполагаем случайные величины x_i $(1 \leqslant i \leqslant n)$ взаимно независимыми.

В силу теоремы о сложении средних значений мы имеем:

$$\overline{X} = \bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}_n$$

и, следовательно,

$$X - \overline{X} = (x_1 - \bar{x}_1) + (x_2 - \bar{x}_2) + \ldots + (x_n - \bar{x}_n),$$

откуда

$$(X - \overline{X})^2 = \left[\sum_{i=1}^{n} (x_i - \bar{x}_i)\right]^2 =$$

$$= \sum_{i=1}^{n} (x_i - \bar{x}_i)^2 + \sum_{\substack{i=1 \\ i \neq k}}^{n} \sum_{k=1}^{n} (x_i - \bar{x}_i)(x_k - \bar{x}_k). \quad (1)$$

Заметим теперь, что

$$\overline{(X - \overline{X})^2} = Q^2, \quad \overline{(x_i - \bar{x}_i)^2} = q_i^2 \quad (1 \leqslant i \leqslant n);$$

пользуясь правилом сложения средних значений, находим поэтому:

$$Q^2 = \sum_{i=1}^{n} q_i^2 + \sum_{\substack{i=1 \\ i \neq k}}^{n} \sum_{k=1}^{n} \overline{(x_i - \bar{x}_i)(x_k - \bar{x}_k)}. \quad (2)$$

Но так как величины x и x_k, согласно нашему предположению, при $i \neq k$ взаимно независимы, то по правилу умножения средних значений взаимно независимых величин при $i \neq k$

$$\overline{(x_i - \bar{x}_i)(x_k - \bar{x}_k)} = \overline{(x_i - \bar{x}_i)}\;\overline{(x_k - \bar{x}_k)}.$$

Здесь оба множителя правой части равны нулю, так как, например,

$$\overline{(x_i - \bar{x}_i)} = \bar{x}_i - \bar{x}_i = 0;$$

таким образом, в последней сумме равенства (2) каждый член в отдельности обращается в нуль, что приводит нас к соотношению

$$Q^2 = \sum_{i=1}^{n} q_i^2,$$

т. е. *дисперсия суммы взаимно независимых случайных величин равна сумме их дисперсий.*

Мы видим, что в случае взаимно независимых случайных величин к правилу сложения средних значений присоединяется весьма важное *правило сложения дисперсий*; для средних квадратических уклонений мы отсюда получаем:

$$Q = \sqrt{\sum_{i=1}^{n} q_i^2}. \qquad (3)$$

Эта возможность просто выразить среднее квадратическое уклонение суммы через средние квадратические уклонения ее слагаемых в случае их взаимной независимости и представляет собой одно из важнейших преимуществ средних квадратических уклонений сравнительно со средними, срединными и другими уклонениями.

Пример 1. Если на некотором предприятии каждое изготовленное изделие может оказаться бракованным с вероятностью p, то среднее число бракованных изделий среди n изготовленных (как мы видели на стр. 85) равно np. Чтобы ориентировочно оценить, сколь большим может оказаться уклонение фактического числа бракованных изделий от этого среднего значения, найдем среднее квадратическое уклонение числа бракованных изделий от np; проще всего это сделать применением формулы (3).

В самом деле, число бракованых изделий можно рассматривать как сумму чисел бракованных изделий при изготовлении каждого изделия (как мы это уже делали при рассмотрении аналогичного примера на стр. 94). А так как эти числа мы считаем взаимно независимыми случайными величинами, то по правилу сложения дисперсий мы можем для вычисления среднего квадратического уклонения Q общего числа бракованных изделий воспользоваться формулой (3), в которой $q_1, q_2, \ldots, q_n$ означают средние квадратические уклонения числа бракованных изделий при изготовлении каждого изделия. Но число бракованных

В частности, если $n=9$, $a_1=a_2=\ldots=a_9=10$ *см* и $q_1=q_2=\ldots=q_9=0{,}2$ *см*, то $a=90$ *см* и $q=\sqrt{9\cdot(0{,}2)^2}=0{,}6$ *см*.

Мы видим, таким образом, что если в среднем длина каждой отдельной детали уклоняется от своего среднего значения на 2%, то длина цепи из этих деталей отличается от своего среднего значения примерно лишь на $^2/_3$%. Это обстоятельство — уменьшение относительной ошибки при сложении случайных величин — играет значительную роль при сборке точных механизмов. Действительно, если бы не было взаимной компенсации отклонений размеров отдельных деталей от заданных нормальных размеров, то при сборке механизмов постоянно встречались бы случаи, когда объемлющие детали не охватывали бы объемлемой цепи или, напротив, при этом оставались бы чрезмерно большие зазоры. В обоих случаях получался бы явный брак; бороться с этим браком по пути уменьшения «допусков», т. е. по пути уменьшения допустимых отклонений фактических размеров детали от заданных, было бы нецелесообразно, так как сравнительно небольшое увеличение точности обработки значительно повышает ее стоимость *).

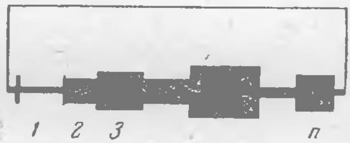

Рис. 10.

Пример 3. Допустим, что в неизменных условиях производятся n измерений некоторой величины. В результате целого ряда обстоятельств (положения прибора, наблюдателя, колебания в состоянии воздуха, наличия в нем пылинок и пр.) различные измерения будут давать, вообще говоря, различные результаты — имеют место *случайные ошибки* измерений. Будем обозначать результаты измерений через $x_1, x_2, \ldots, x_n$,

*) Техническая мысль в последнее время пришла к выводу о необходимости создания теории допусков, основанной на соображениях и выводах теории вероятностей. Эта теория допусков сейчас усиленно разрабатывается учеными.

изделий при изготовлении i-го из них определяется таблицей

1	0
p	$1-p$

поэтому $\bar{x}_i = p$ и

$$q_i^2 = \overline{(x_i - \bar{x}_i)^2} = (1-p)^2 p + p^2 (1-p) = p(1-p);$$

следовательно,

$$Q = \sqrt{\sum_{i=1}^{n} q_i^2} = \sqrt{np(1-p)}.$$

Этим поставленная задача решена.

Сопоставляя среднее число np бракованных изделий с его средним квадратическим уклонением $\sqrt{np(1-p)}$, мы видим, что при больших значениях n последнее значительно меньше первого и составляет лишь небольшую долю его. Так, при $n = 60\,000$, $p = 0{,}04$ среднее число бракованных изделий равно 2400, а среднее квадратическое уклонение $Q = \sqrt{60\,000 \cdot 0{,}04 \cdot 0{,}96} = 48$, так что фактическое число бракованных изделий ориентировочно будет уклоняться лишь на 5% от своего среднего значения.

Пример 2. Представим себе, что производится сборка некоторого механизма, состоящего из n деталей, прикладываемых вплотную одна к другой вдоль некоторой оси и охватываемых с концов некоторой объемлющей деталью (рис. 10). Длина каждой детали может несколько отличаться от соответствующего стандарта и есть поэтому случайная величина. Предположим эти случайные величины независимыми. Если средние длины деталей и дисперсии этих длин равны соответственно

$$a_1,\ a_2,\ \ldots,\ a_n \text{ и } q_1,\ q_2,\ \ldots,\ q_n,$$

то среднее значение и дисперсия длины цепи из n деталей равны

$$a = \sum_{k=1}^{n} a_k \quad \text{и} \quad q = \sqrt{\sum_{k=1}^{n} q_k^2}\,.$$

приписывая каждому x в качестве индекса (значка) номер измерения. Среднее значение для всех этих случайных величин одно и то же — $\bar{x}$. Среднее квадратическое уклонение q, очевидно, также естественно предположить одним и тем же для всех измерений, так как они производятся в неизменных условиях. Наконец, мы считаем, как обычно, величины $x_1, x_2, \ldots, x_n$ взаимно независимыми.

Рассмотрим теперь среднее арифметическое

$$\xi = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

результатов n измерений. Это — случайная величина; найдем ее среднее значение и среднее квадратическое уклонение. По правилу сложения

$$\bar{\xi} = \frac{1}{n}\overline{(x_1 + x_2 + \ldots + x_n)} = \frac{1}{n}(\bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}_n) = \\ = \frac{1}{n}(n\bar{x}) = \bar{x},$$

т. е. среднее значение, как это в сущности и ранее было очевидно, то же самое, что и для каждого отдельного измерения.

Далее, среднее квадратическое уклонение суммы $x_1 + x_2 + \ldots + x_n$ по правилу сложения дисперсий (3) равно

$$Q = \sqrt{nq^2} = q\sqrt{n},$$

а среднее квадратическое уклонение величины ξ, составляющей $^1/_n$ этой суммы, равно

$$\frac{Q}{n} = \frac{q}{\sqrt{n}}.$$

Мы приходим к очень важному результату: *среднее арифметическое n взаимно независимых и одинаково распределенных случайных величин имеет:*

а) *среднее значение — то же самое, что и каждая из составляющих величин;*

б) *среднее квадратическое уклонение — в $\sqrt{n}$ раз меньшее, чем каждая из составляющих величин.*

Так, если среднее значение измеряемой величины $x = 200$ *м*, а среднее квадратическое уклонение $q = 5$ *м*, то среднее арифметическое ξ сотни результатов измерений будет, конечно, иметь своим средним значением то же число 200 *м*; но среднее квадратическое уклонение его будет в $\sqrt{100} = 10$ раз меньше, чем для отдельного измерения, т. е. будет составлять всего $\frac{q}{10} = 0{,}5$ *м*. Таким образом, имеются основания ожидать, что среднее арифметическое сотни фактических результатов измерений будет значительно ближе к среднему значению 200 *м*, чем результат того или другого отдельного измерения. *Среднее арифметическое большого числа взаимно независимых величин обладает во много раз меньшим рассеянием, чем каждая из этих величин в отдельности.*

ГЛАВА ОДИННАДЦАТАЯ

ЗАКОН БОЛЬШИХ ЧИСЕЛ

§ 26. Неравенство Чебышева

Мы говорили уже много раз о том, что знание какого-либо из средних уклонений случайной величины (например, ее среднего квадратического уклонения) позволяет нам создать ориентировочное представление о том, как велики те уклонения фактических значений этой величины от ее среднего значения, которые нам следует ожидать. Однако это замечание само по себе не содержит еще никаких количественных оценок и не дает возможности хотя бы приближенно рассчитать, сколь вероятными могут оказаться большие уклонения. Все это позволяет сделать следующее простое рассуждение, проведенное впервые Чебышевым. Будем исходить из выражения дисперсии случайной величины x (стр. 103):

$$Q_x^2 = \sum_{i=1}^{k} (x_i - \bar{x})^2 p_i.$$

Пусть α — любое положительное число; если в только что написанной сумме мы выбросим все члены, где $|x_i - \bar{x}| \leqslant \alpha$, и оставим только те, где $|x_i - \bar{x}| > \alpha$, то от этого сумма может только уменьшиться:

$$Q_x^2 \geqslant \sum_{|x_i - \bar{x}| > \alpha} (x_i - \bar{x})^2 p_i.$$

Но эта сумма уменьшится еще более, если в каждом члене мы заменим множитель $(x_i - \bar{x})^2$ меньшей величиной α^2:

$$Q_x^2 \geqslant \alpha^2 \sum_{|\bar{x}_i - \bar{x}| > \alpha} p_i.$$

Сумма, стоящая теперь в правой части, есть сумма вероятностей всех тех значений x_i случайной величины x, которые уклоняются от $\bar{x}$ в ту или другую сторону больше, чем на α; по правилу сложения это есть вероятность того, что величина x получит какое-либо одно из этих значений. Другими словами, это есть вероятность $\mathsf{P}(|x - \bar{x}| > \alpha)$ того, что фактически полученное уклонение окажется больше, чем α; таким образом, мы находим

$$\mathsf{P}(|x - \bar{x}| > \alpha) \leqslant \frac{Q_x^2}{\alpha^2}. \tag{1}$$

Полученное соотношение называется *неравенством Чебышева*. Оно позволяет нам оценить вероятность уклонений, больших чем любое заданное число α, если только известно среднее квадратическое уклонение Q_x. Правда, оценка, даваемая неравенством Чебышева, часто оказывается весьма грубой; все же иногда она может быть использована практически не говоря уже о том, что теоретическое значение ее чрезвычайно велико.

В конце предыдущего параграфа мы рассматривали такой пример: среднее значение результатов измерений — 200 *м*, среднее квадратическое уклонение — 5 *м*; при этих условиях вероятность фактически получить уклонение больше 3 *м* весьма ощутительна (можно думать, что она больше половины; точное значение ее может, конечно, быть найдено только тогда, когда полностью известен закон распределения результатов измерений). Но мы видели, что для среднего арифметического ξ сотни результатов измерений среднее квадратическое уклонение составляет всего 0,5 *м*. Поэтому в силу неравенства (1)

$$\mathsf{P}(|\xi - 200| > 3) \leqslant \frac{(0{,}5)^2}{3^2} = \frac{1}{36} \approx 0{,}03.$$

Таким образом, для среднего арифметического из 100 измерений вероятность получить уклонение более 3 *м* уже очень мала (на самом деле она еще значительно меньше полученной нами границы, так что практически можно совсем не считаться с возможностью такого уклонения).

В примере 1 на стр. 109—110 для числа бракованных изделий при проверке 60 000 изделий мы получили среднее значение 2400 и среднее квадратическое уклонение 48. Для вероятности того, что фактическое число бракованных изделий будет заключено, например, между 2300 и 2500, т. е. $|m-2400| \leqslant 100$, неравенство Чебышева дает

$$\mathrm{P}\{|m-2400| \leqslant 100\} = 1 - \mathrm{P}\{|m-2400| > 100\} \geqslant$$
$$\geqslant 1 - \frac{48^2}{100^2} \approx 0{,}77.$$

На самом деле эта вероятность значительно больше.

§ 27. Закон больших чисел

Пусть мы имеем n взаимно независимых случайных величин $x_1, x_2, \ldots, x_n$ с одним и тем же средним значением a и одним и тем же средним квадратическим уклонением q; для среднего арифметического этих величин

$$\xi = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

как мы видели на стр. 112 среднее значение равно a, а среднее квадратическое уклонение равно $\frac{q}{\sqrt{n}}$; поэтому неравенство Чебышева дает при любом положительном α

$$\mathrm{P}(|\xi - a| > \alpha) \leqslant \frac{q^2}{\alpha^2 n}. \qquad (2)$$

Пусть, например, речь идет о среднем арифметическом n измерений некоторой величины, и пусть, как мы имели прежде, $q = 5$ *м*, $a = 200$ *м*. Тогда мы получаем:

$$\mathrm{P}(|\xi - 200| > \alpha) \leqslant \frac{25}{\alpha^2 n}.$$

Мы можем выбрать α очень небольшим, например $\alpha = 0{,}5$ *м*; тогда

$$\mathsf{P}(|\xi - 200| > 0{,}5) \leqslant \frac{100}{n}.$$

Если число измерений n очень велико, то правая часть этого неравенства как угодно мала; так, при $n = 10\,000$ она равна 0,01, и мы имеем для среднего арифметического 10 000 измерений

$$\mathsf{P}(|\xi - 200| > 0{,}5) \leqslant 0{,}01.$$

Если условиться пренебрегать возможностью столь маловероятных событий, то можно сказать, что при 10 000 измерений их среднее арифметическое наверняка будет отличаться от 200 *м* в ту или в другую сторону не более чем на 50 *см*. Если бы мы захотели достигнуть еще большей близости, например 10 *см*, то надо было бы положить $\alpha = 0{,}1$ *м*, и тогда

$$\mathsf{P}(|\xi - 200| > 0{,}1) \leqslant \frac{25}{0{,}01n} = \frac{2500}{n}.$$

Чтобы сделать менее 0,01 правую часть этого неравенства, мы должны были бы взять число измерений равным не 10 000 (этого теперь недостаточно), а 250 000. Очевидно, что вообще можно сколь бы мало ни было α, сделать правую часть неравенства (2) как угодно малой — для этого стоит только взять n достаточно большим. Таким образом, при достаточно большом n можно считать сколь угодно близким к достоверности обратное неравенство

$$|\xi - a| \leqslant \alpha.$$

Если случайные величины $x_1, x_2, \ldots, x_n$ взаимно независимы и если все они имеют одно и то же среднее значение a и одно и то же среднее квадратическое уклонение, то величина

$$\xi = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

при достаточно большом n будет с вероятностью, как угодно близкой к единице (практически достоверно), как угодно мало отличаться от a.

Это — простейший частный случай одной из самых основных теорем теории вероятностей, так называемого *закона больших чисел*, открытого в середине прошлого столетия великим русским математиком Чебышевым. Глубокое содержание этого замечательного закона состоит в том, что, в то время как отдельная случайная величина может (как мы знаем) часто принимать значения, далекие от ее среднего значения (иметь значительное рассеяние), среднее арифметическое большого числа случайных величин ведет себя в этом отношении совершенно иначе: такая величина очень мало рассеяна, с подавляющей вероятностью она принимает лишь значения, очень близкие к ее среднему значению. Происходит это, конечно, потому, что при взятии среднего арифметического случайные уклонения в ту и другую сторону взаимно уничтожаются, вследствие чего суммарное уклонение в большинстве случаев и оказывается малым.

Важное и часто встречающееся в практике использование результатов только что доказанной теоремы Чебышева состоит в том, что по сравнительно небольшой пробе (выборке) судят о качестве однородного материала. Так, например, о качестве хлопка, находящегося в кипе, судят по нескольким маленьким его пучочкам (штапелям), выхваченным случайно из разных мест кипы. Или о качестве большой партии зерна судят по нескольким небольшим пуркам (меркам), наполненным случайно захваченными в пурку зернами из разных мест оцениваемой партии *). Суждения о качестве продукции, сделанные на основании такой выборки, обладают большой точностью, так как, скажем, число зерен, захваченных в пурку, хотя и мало по сравнению со всем запасом зерна, но само по себе велико и позволяет, согласно закону больших чисел, достаточно точно судить о среднем весе одного зерна и, значит, о качестве всей партии зерна. Точно так же

*) В пурку захватывают, скажем, 100—200 граммов, а вся партия содержит десятки, а может быть даже сотни тонн зерна.

о двадцатипудовой кипе хлопка судят по маленькому штапелю, содержащему несколько сотен волокон, весящих всего-навсего какую-нибудь десятую долю грамма.

§ 28. Доказательство закона больших чисел

До сих пор мы рассматривали только случай, когда все величины $x_1, x_2, \ldots$ имеют одно и то же среднее значение и одно и то же среднее квадратическое уклонение. Однако закон больших чисел применяется и в более общих предположениях. Будем теперь рассматривать случай, когда средние значения величин $x_1, x_2, \ldots$ могут быть какими угодно числами (обозначим их соответственно через $a_1, a_2, \ldots$), вообще говоря, между собою различными. Тогда средним значением величины

$$\xi = \frac{1}{n}(x_1 + x_2 + \ldots + x_n),$$

очевидно, будет величина

$$A = \frac{1}{n}(a_1 + a_2 + \ldots + a_n),$$

причем, в силу неравенства Чебышева (1), при любом положительном α

$$\mathsf{P}(|\xi - A| > \alpha) \leqslant \frac{Q_\xi^2}{\alpha^2}. \qquad (3)$$

Мы видим, что все сводится к оценке величины Q_ξ^2, но оценить эту величину здесь почти так же просто, как и в ранее рассмотренном частном случае. Q_ξ^2 есть дисперсия величины ξ, которая равна разделенной на n сумме n взаимно независимых случайных величин (предположение о взаимной независимости, разумеется, сохраняем). По правилу сложения дисперсий имеем поэтому

$$Q_\xi^2 = \frac{1}{n^2}(q_1^2 + q_2^2 + \ldots + q_n^2),$$

где $q_1, q_2, \ldots$ соответственно означают средние квадратические уклонения величин $x_1, x_2, \ldots$ Теперь будем предполагать, что эти средние квадратические уклонения, вообще говоря, также различны между собой. Однако допустим все же, что, сколь бы много величин мы ни брали (т. е. сколь бы велико ни было число n), средние квадратические уклонения всех этих величин всегда будут меньше некоторого положительного числа b. На практике это условие всегда оказывается выполненным, так как складывать приходится величины более или менее однотипные, и степень их рассеяния оказывается не слишком различной для различных величин.

Итак, допустим, что $q_i < b$ $(i = 1, 2, \ldots)$; но тогда последнее равенство дает

$$Q_\xi < \frac{1}{n^2} nb^2 = \frac{b^2}{n},$$

вследствие чего заключаем из неравенства (3), что

$$\mathrm{P}(|\xi - A| > \alpha) < \frac{b^2}{n\alpha^2}.$$

Как бы мало ни было α, при достаточно большом числе n взятых случайных величин правая часть этого неравенства может быть сделана как угодно малой; это, очевидно, и доказывает закон больших чисел в рассматриваемом теперь общем случае.

Итак, *если величины $x_1, x_2, \ldots$ взаимно независимы, а средние квадратические уклонения их все остаются меньше одного и того же положительного числа, то при достаточно больших n для среднего арифметического*

$$\xi = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

можно с вероятностью, сколь угодно близкой к единице, ожидать сколь угодно малых по абсолютной величине уклонений.

Это и есть закон больших чисел в общей форме, приданной ему Чебышевым.

Сейчас уместно обратить внимание на одно существенное обстоятельство. Предположим, что производится измерение некоторой величины a. Повторив измерения в одинаковых условиях, наблюдатель получит не вполне совпадающие числовые результаты x_1, x_2, ..., x_n. В качестве приближенного значения a он берет среднюю арифметическую

$$a \sim \frac{1}{n}(x_1 + x_2 + \ldots + x_n).$$

Спрашивается, можно ли рассчитывать получить сколь угодно точное значение a, если произвести достаточно большое число опытов?

Это будет так, если измерения производятся без систематической ошибки, т. е. если

$$\bar{x}_k = a \qquad (\text{при } k = 1, 2, \ldots, n)$$

и сами значения не обладают неопределенностью, иными словами, если при измерениях мы читаем те показания на приборе, которые там в действительности получаются. Если же прибор устроен так, что он не может давать точности отсчета, большей чем некоторая величина δ, например из-за того, что ширина деления шкалы, по которой производится отсчет, равна δ, то понятно, что нельзя и рассчитывать получить точность, большую чем $\pm\delta$. Ясно, что в этом случае и средняя арифметическая будет обладать той же неопределенностью δ, как и каждое из x_k.

Сделанное замечание учит нас, что если приборы дают нам результаты измерений с некоторой неопределенностью δ, то стремиться посредством закона больших чисел получить значение a с большей точностью является заблуждением, а сами производимые при этом вычисления превращаются в пустую арифметическую забаву.

ГЛАВА ДВЕНАДЦАТАЯ

НОРМАЛЬНЫЕ ЗАКОНЫ

§ 29. Постановка задачи

Мы видели, что значительное количество явлений природы, а также производственных процессов и операций протекает при существенном участии тех или иных случайных величин. Часто до того, как явление, процесс или операция не завершены, все, что мы можем знать об этих случайных величинах, — это их законы распределения, т. е. списки их возможных значений с указанием вероятности каждого из этих значений. Если величина может получать бесчисленное множество различных значений (дальность полета снаряда, величина ошибки измерения и т. п.), то предпочтительнее указывать вероятности не отдельных значений ее, а целых участков таких значений (например, вероятность того, что ошибка измерения будет заключена в пределах от —1 *мм* до +1 *мм*, от 0,1 *мм* до 0,25 *мм* и т. д.).

Это соображение не меняет существа дела, — чтобы стать хозяевами случайной величины, чтобы овладеть ею в меру наших сил и возможностей, необходимо получить возможно точное представление о ее законе распределения.

Если бы мы, стараясь узнать законы распределения встречающихся случайных величин, отказались от всяких рассуждений и догадок общего характера и к каждой случайной величине подходили без всяких предварительных предположений, стремясь найти чисто опытным путем все черты свойственного ей закона распределения, то поставили бы себя перед задачей, практически почти невыполнимой по своей трудоемкости. В каждом новом случае потребовалось бы большое число опытов, чтобы установить хотя бы важнейшие черты нового, неизвестного нам закона распределения. Поэтому ученые уже давно старались найти такие общие типы законов распределения, наличие которых можно было бы с разумным основанием предвидеть, ожидать, подозревать, если не для всех, то по крайней мере для широких классов встречающихся на практике случайных величин. И уже давно некоторые такие типы были теоретически установлены, а затем наличие их было подтверждено опытом. Понятно, насколько выгодно бывает, на основании теоретических рассуждений и всего предшествующего опыта, заранее иметь возможность предугадать, какого типа должен быть закон распределения вновь встретившейся нам случайной величины. Если такая догадка оказывается правильной, то обычно уже весьма небольшого числа опытов или наблюдений бывает достаточно, чтобы установить все нужные нам черты искомого закона распределения.

Теоретические исследования показали, что в большом числе встречающихся на практике случаев мы с достаточным основанием можем ожидать законов распределения некоторого совершенно определенного типа. Эти законы называются *нормальными законами.* О них мы кратко, опуская ввиду сложности все доказательства и точные формулировки, расскажем в настоящей главе.

Среди случайных величин, встречающихся нам в практике, очень многие носят характер «случайных погрешностей», или «случайных ошибок», или по крайней мере легко сводятся к таким «погрешностям». Пусть, например, изучается дальность x полета снаряда при стрельбе из некоторого орудия. Мы, естественно, допускаем, что существует некоторая нормальная, средняя дальность x_0, на которую и устанавливаем прицельные приборы; разность $x - x_0$ составляет «погрешность», или «ошибку», дальности, и изучение случайной величины x целиком и непосредственно сводится к изучению «случайной ошибки» $x - x_0$.

Но такая ошибка, меняющая свою величину от выстрела к выстрелу, зависит, как правило, от очень многих причин, действующих независимо друг от друга: случайные колебания ствола орудия, неизбежный (хотя бы и небольшой) разнобой в весе и форме снарядов, случайные изменения атмосферных условий, вызывающие изменения в сопротивлении воздуха, случайные ошибки в наводке (если наводка производится заново перед каждым выстрелом или каждой небольшой группой выстрелов), — все эти и многие другие причины способны вызвать ошибки в дальности полета. Все частичные ошибки будут независимыми между собой случайными величинами, причем такими, что *действие каждой из них составляет лишь очень малую долю их совокупного действия*, а окончательная ошибка $x - x_0$, которую мы хотим исследовать, будет просто суммарным действием всех случайных ошибок, происходящих от отдельных причин. Таким образом, в данном примере интересующая нас ошибка является суммой большого числа взаимно независимых случайных величин, и ясно, что подобным же образом дело будет обстоять для большинства случайных ошибок, с которыми мы имеем дело на практике.

И вот, теоретическое рассуждение, которого мы здесь не может воспроизвести, показывает, что закон распределения случайной величины, являющейся суммой очень большого числа взаимно независимых случайных величин, уже в силу одного этого, какова бы ни была природа слагаемых, *лишь бы каждое из них было мало по сравнению со всей суммой*, должен быть близок к закону некоторого совершенно определенного типа *).

Этот тип и есть тип нормальных законов. Таким образом, мы получаем возможность предполагать, что весьма значительная часть встречающихся в практике случайных величин (все ошибки, складывающиеся из

---

*) См. об этом также заключение.

большого числа взаимно независимых случайных ошибок) приблизительно распределена по нормальным законам. Теперь мы должны будем ознакомиться с основными чертами этих законов.

§ 30. Понятие кривой распределения

В § 15 мы уже имели случай изображать законы распределения графически, с помощью диаграмм; этот способ очень полезен, так как позволяет одним взглядом, не прибегая к исследованию таблиц, охватить

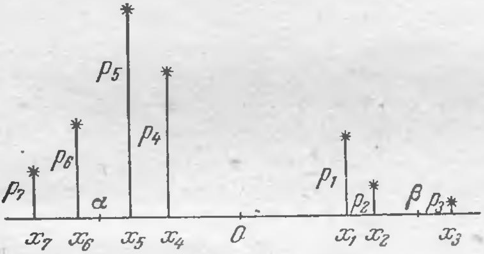

Рис. 11.

важнейшие черты исследуемого закона распределения. Используемая схема изображения такова: на горизонтальной прямой откладываются различные возможные значения данной случайной величины, начиная от некоторого начала отсчета O, положительные вправо, отрицательные влево (рис. 11). Против каждого возможного значения откладывают по вертикали кверху вероятность этого значения. Масштаб в обоих направлениях выбирается такой, чтобы вся диаграмма имела удобную и легко обозримую форму. Один беглый взгляд на рис. 11 убеждает нас в том, что характеризуемая им случайная величина имеет наивероятнейшее значение x_5 (отрицательное) и что по мере удаления возможных значений этой величины от числа x_5 вероятности их непрестанно (и довольно быстро) убывают. Вероятность того, что величина примет значение, заключенное в каком-либо отрезке (α, β), по правилу сложения равна сумме вероятностей всех возможных значений, лежащих в этом отрезке, и геометрически изображается суммой длин вертикальных черточек, расположенных над этим отрезком; на рис. 11 $\mathrm{P}(\alpha < x < \beta) = p_1 + p_2 + p_4 + p_5$. Если, как это часто бывает на практике, число возможных значений очень велико, то, чтобы чертеж не слишком вытянулся по горизонтали, берут очень малый масштаб в горизонтальном направлении, вследствие чего возможные значения располагаются чрезвычайно густо (рис. 12), так

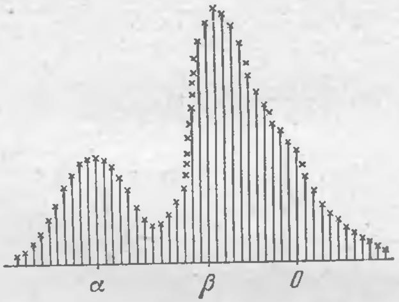

Рис. 12.

что верхушки проведенных вертикальных черточек сливаются для нашего глаза в одну сплошную кривую линию, которую называют *кривой распределения* данной случайной величины. И здесь, конечно, вероятность неравенств $\alpha < x < \beta$ графически изображается суммой длин вертикальных черточек, расположенных над отрезком (α, β).

Допустим теперь, что расстояние между двумя соседними возможными значениями всегда равно единице; это будет, например, если возможные значения выражаются рядом последовательных целых чисел, чего практически всегда можно достигнуть, выбирая достаточно мелкую единицу масштаба. Тогда длина каждой вертикальной черточки численно равна площади прямоугольника, высотой которого служит эта черточка, а основанием — равное единице расстояние ее от соседней черточки (рис. 13). Таким образом, вероятность неравенств $\alpha < x < \beta$ графически может быть изображена суммой площадей нарисованных на чертеже прямоугольников, расположенных над этим отрезком. Но если возможные значения расположены очень густо, как на рис. 12, то сумма площадей таких прямоугольников практически не будет отличаться от площади криволинейной фигуры, ограниченной снизу отрезком (α, β), сверху — кривой распределения, а с боков — вертикальными черточками, проведенными из точек α и β (рис. 14) *). Таким образом,

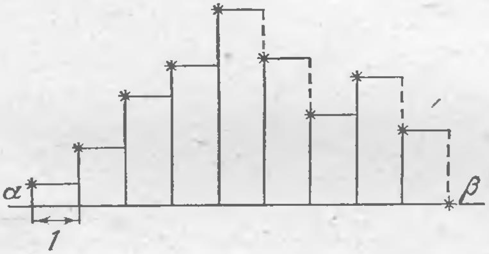

Рис. 13.

на криволинейной диаграмме типа рис. 14 вероятность попадания данной случайной величины в любой отрезок просто и удобно выражается площадью, лежащей над этим отрезком ниже кривой распределения. Если закон распределения данной величины задается такой криволинейной диаграммой, то на ней вовсе не проводят вертикальных черточек, которые только загромоздили бы собой чертеж; да и самый вопрос о вероятностях отдельных значений здесь теряет свою актуальность; если возможных значений очень много (а ведь именно это и лежит в основании всех криволинейных диаграмм), то вероятности отдельных значений будут, как правило, ничтожно малы (практически равны нулю) и теряют всякий интерес. Так, при измерении расстояния между двумя населенными пунктами совсем несущественно знать, что результат измерения отклонится от истинного значения ровно на 473 *см*.

*) При этом, конечно, по-прежнему за единицу длины принимается расстояние между двумя соседними возможными значениями.

Напротив, представляет существенный интерес вопрос о вероятности уклонения, заключенного в промежутке от 3 *м* до 5 *м*. И так во всех подобных случаях: если случайная величина принимает очень много значений,

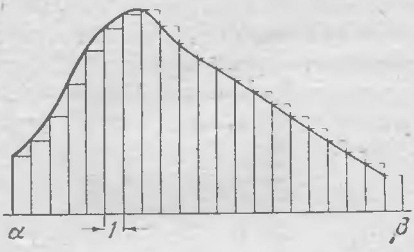

Рис. 14.

то нам важно знать вероятности не отдельных этих значений, а вероятности целых отрезков таких значений. Но именно эти вероятности даются наглядно и непосредственно, как мы только что видели, площадями на криволинейных диаграммах.

§ 31. Свойства нормальных кривых распределения

Величина, распределенная по нормальному закону, всегда имеет бесчисленное множество возможных значений; поэтому нормальные законы удобно графически изображать криволинейными диаграммами.

На рис. 15 изображено несколько кривых распределения по нормальному закону. Несмотря на все различия в облике этих кривых, мы видим у них ясно выраженные общие им всем черты:

1) Все кривые имеют одну наивысшую точку, при удалении от которой вправо или влево они непрестанно понижаются. Очевидно, это означает, что при удалении значений случайной величины от ее наивероятнейшего значения вероятности их непрестанно убывают.

2) Все кривые симметричны относительно вертикальной прямой, проведенной через наивысшую точку. Очевидно, это означает, что значения, равноудаленные от наивероятнейшего значения, имеют одинаковые вероятности.

3) Все кривые имеют колоколообразную форму; вблизи наивысшей точки они обращены выпуклостью кверху, а на некотором расстоянии от нее перегибаются и обращаются выпуклостью книзу; расстояние это (как и наибольшая высота) различно для различных кривых *).

Чем же отличаются друг от друга различные нормальные кривые? Чтобы ясно ответить на этот вопрос, мы должны прежде всего вспомнить, что для всякой кривой распределения вся расположенная под нею площадь равна единице, ибо площадь эта равна вероятности того, что данная случайная величина примет какое бы-то ни было из своих значений, т. е. вероятности достоверного события. Отличие отдельных кривых распределения друг от друга состоит поэтому лишь в том, что эта суммарная площадь, одна и та же для всех кривых, различным образом распределена между различными участками. Для нормальных законов, как показывают кривые на рис. 15, вопрос в основном заключается в том, какая доля этой суммарной площади сосредоточена над участками, непосредственно примыкающими к наивероятнейшему значению, и какая — над участками, более удаленными от этого значения. Для закона, изображаемого рис. 15, *а*), почти вся площадь сосредоточена в непосредственной близости наивероятнейшего значения; это означает, что случайная величина с подавляющей вероятностью — и, значит, в подавляющем большинстве случаев — принимает значения, близкие к ее наивероят-

*) Для читателей, знакомых с элементами высшей математики, мы заметим, что уравнение кривой, изображающей нормальный закон, имеет такой вид:

$$y = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right),$$

где число $e = 2{,}71828\ldots$ есть основание натуральных логарифмов, $\pi = 3{,}14159\ldots$ — отношение длины окружности к ее диаметру, а величины a и σ^2 представляют собой среднее значение и дисперсию случайной величины. Знание аналитической формы нормального закона может значительно облегчить читателю ознакомление с дальнейшим материалом книги. Однако изложение всего последующего доступно и читателю, не знакомому с высшей математикой.

нейшему значению; так как в силу указанной выше симметрии в случае нормального закона наивероятнейшее значение всегда совпадает со средним значением, то мы можем сказать, что случайная величина, подчиненная закону *а*), мало рассеяна; в частности, ее дисперсия и среднее квадратическое уклонение малы.

Наоборот, в случае, изображенном на рис. 15, *в*), площадь, сосредоточенная в непосредственной близости наивероятнейшего значения, составляет лишь небольшую долю суммарной площади (мы сразу увидим

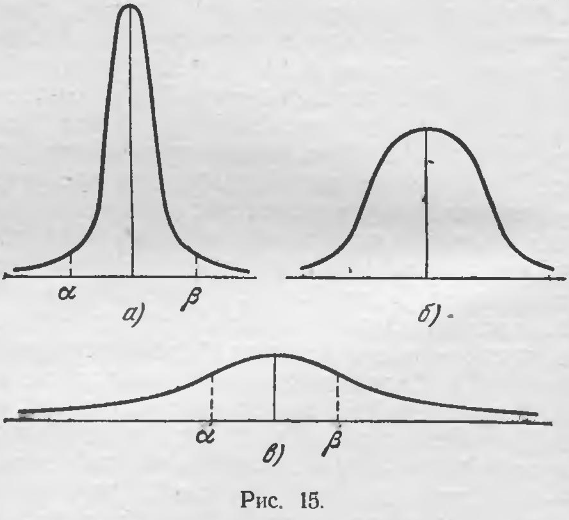

Рис. 15.

различие, если сопоставим на рис. 15, *а*) и *в*) участки (α, β) одинаковой длины и расположенные над ними площади). Здесь весьма вероятно поэтому, что случайная величина будет получать значения, заметно уклоняющиеся от ее наивероятнейшего значения. Величина сильно рассеяна, ее дисперсия и среднее квадратическое уклонение велики.

Случай *б*), очевидно, занимает положение, промежуточное между случаями *а*) и *в*).

Чтобы наиболее быстрым образом ознакомиться со всей совокупностью нормальных законов и научиться применять их, целесообразно исходить из двух основных свойств, которые будут сейчас подробно сформулированы, но не доказаны, так как для этого нам пришлось бы прежде всего точно определить нормальные законы, что потребовало бы от читателя знания высшей математики.

Свойство 1. *Если величина x распределена по нормальному закону*, то

1) *при любых постоянных $c > 0$ и d величина $cx + d$ также распределена по некоторому нормальному закону, и*

2) *обратно, для любого нормального закона найдется такая (единственная) пара чисел $c > 0$ и d, что величина $cx + d$ распределена именно по этому закону.*

Таким образом, если случайная величина x распределена по нормальному закону, то законы распределения, которым подчиняются величины $cx + d$ при всевозможных значениях постоянных $c > 0$ и d, — это все нормальные законы.

Свойство 2. *Если случайные величины x и y взаимно независимы и распределены по нормальным законам, то и сумма их $z = x + y$ распределена по некоторому нормальному закону.*

Приняв без доказательства эти два основных свойства, мы можем теперь строго обосновать ряд свойств нормальных законов, особенно важных для практики.

I. *Для любых двух чисел a и $q > 0$ существует единственный нормальный закон со средним значением a и средним квадратическим уклонением q.*

В самом деле, пусть x — случайная величина, распределенная по нормальному закону со средним значением $\bar{x}$ и средним квадратическим уклонением Q_x. На основании свойства 1 наше утверждение будет доказано, если мы покажем, что существует единственная пара чисел $c > 0$ и d, удовлетворяющая тому требованию, чтобы величина $cx + d$ имела среднее значение a и среднее квадратическое уклонение q. Если таблица величины x имеет вид

x_1	x_2	...	x_n
p_1	p_2	...	p_n

то величине $cx+d$ (где $c>0$ и d — пока любые постоянные) будет соответствовать таблица

cx_1+d	cx_2+d	...	cx_n+d
p_1	p_2	...	p_n

Очевидно, $\sum\limits_k x_k p_k = \bar{x}$, $\sum\limits_k (x_k-\bar{x})^2 p_k = Q_x^2$ *).

Наши требования сводятся к двум условиям:

$$\sum_k (cx_k+d)\,p_k = a; \quad \sum_k (cx_k+d-a)^2 p_k = q^2.$$

Первое из этих условий дает $c\sum\limits_k x_k p_k + d\sum\limits_k p_k = a$, или

$$c\bar{x}+d=a, \tag{1}$$

а второе $\sum\limits_k (cx_k+d-c\bar{x}-d)^2 p_k = c^2 \sum\limits_k (x_k-\bar{x})^2 p_k = c^2Q_x^2 = q^2$, откуда (так как $c>0$)

$$c=\frac{q}{Q_x}, \tag{2}$$

и, значит, из (1)

$$d=a-c\bar{x}=a-\frac{q\bar{x}}{Q_x}. \tag{3}$$

Таким образом, по данным a и q числа c и d с помощью (2) и (3) всегда могут быть найдены, и притом единственным способом; величина $cx+d$ подчиняется нормальному закону со средним значением a и средним квадратическим уклонением q; этим наше утверждение доказано.

Если не ограничиваться нормальными законами, а рассматривать всевозможные законы распределения, то задание среднего значения и дисперсии или среднего квадратического уклонения случайной величины дает нам об ее законе распределения еще очень мало сведений, так как существует весьма много (и притом

*) Знак $\sum\limits_k$ — сокращенный; он обозначает $\sum\limits_{k=1}^{n}$.

существенно различных между собой) законов распределения, обладающих одним и тем же средним значением и одной и той же дисперсией; в общем случае задание среднего значения и дисперсии лишь весьма приблизительно характеризует закон распределения данной случайной величины.

Иначе обстоит дело, если ограничиться рассмотрением одних только нормальных законов. С одной стороны, как мы только что убедились, любое предположение о среднем значении и дисперсии данной случайной величины совместимо с требованием, чтобы она подчинялась нормальному закону. С другой стороны, — и это самое главное, — если мы имеем основание заранее предполагать, что данная величина распределена по одному из нормальных законов, то заданием ее среднего значения и дисперсии закон распределения однозначно определяется, так что ее природа как случайной величины становится полностью известной. В частности, зная среднее значение и дисперсию такой величины, мы можем вычислить вероятность того, что значение ее будет принадлежать тому или другому произвольно выбранному участку.

II. *Отношение срединного (вероятного) уклонения к среднему квадратическому уклонению одно и то же для всех нормальных законов.*

Пусть мы имеем два произвольных нормальных закона, и пусть x — случайная величина, подчиняющаяся первому из этих законов. В силу основного свойства 1 существуют такие постоянные числа $c > 0$ и d, что величина $cx + d$ распределена согласно второму из данных законов. Обозначим соответственно через Q_x и E_x среднее квадратическое уклонение и срединное (вероятное) уклонение первой величины и через q и e — те же уклонения для второй величины. По определению вероятного уклонения

$$\mathsf{P}\{|(cx+d)-(c\bar{x}+d)|<e\}=\frac{1}{2},$$

или

$$\mathsf{P}\{c\,|x-\bar{x}|<e\}=\frac{1}{2},$$

или, наконец,

$$\mathsf{P}\left(|x-\bar{x}|<\frac{e}{c}\right)=\frac{1}{2}.$$

Отсюда, снова по определению вероятного уклонения, следует, что $\frac{e}{c}$ есть вероятное уклонение величины x, т. е. $\frac{e}{c}=E_x$, откуда $\frac{e}{E_x}=c$; поэтому из (2) $\frac{e}{E_x}=\frac{q}{Q_x}$, откуда $\frac{e}{q}=\frac{E_x}{Q_x}$, т. е. отношение вероятного уклонения к среднему квадратическому уклонению одинаково для наших двух законов.

Так как эти законы, по предположению, были двумя произвольными нормальными законами, то наше утверждение доказано.

Отношение $\frac{e}{q}$ есть, таким образом, абсолютная постоянная; обозначим ее λ; вычислено, что $\lambda=\sqrt{\frac{2}{\pi}}\approx$ $\approx 0,674$. Значит, для любого нормального закона $e=\sqrt{\frac{2}{\pi}}q$.

В силу этой исключительно простой связи между числами e и q для величин, распределенных по нормальным законам, практически безразлично, какой из двух характеристик рассеяния пользоваться; выше говорилось, что (даже если не ограничиваться величинами, распределенными по нормальным законам) среднее квадратическое уклонение обладает целым рядом простых свойств, которых лишены другие характеристики и которые в большинстве случаев заставляют как теоретиков, так и практиков выбирать в качестве меры рассеяния именно средние квадратические уклонения. Там же было указано, что артиллеристы тем не менее почти всегда пользуются срединными уклонениями. Мы видим теперь, почему эта традиция не может принести никакого ущерба: случайные величины, с которыми имеют дело артиллерийская наука и практика, почти всегда оказываются распределенными по нормальным законам, а для таких величин, в силу упомянутой выше пропорциональности, выбор любой из двух характеристик практически безразличен.

III. *Пусть x и y — вазимно независимые случайные величины, подчиненные нормальным законам, и $z = x + y$. Тогда*

$$E_z = \sqrt{E_x^2 + E_y^2},$$

где E_x, E_y, E_z означают соответственно вероятные уклонения величин x, y, z.

Аналогичная формула для средних квадратических уклонений, как мы знаем из § 25, имеет место, каковы бы ни были законы распределения величин x и y. В случае, когда это нормальные законы, величина z в силу основного свойства 2 также распределена по нормальному закону; поэтому на основании свойства II

$$E_x = \lambda Q_x, \quad E_y = \lambda Q_y, \quad E_z = \lambda Q_z,$$

и, значит,

$$E_z = \lambda \sqrt{Q_x^2 + Q_y^2} = \sqrt{(\lambda Q_x)^2 + (\lambda Q_y)^2} = \sqrt{E_x^2 + E_y^2}.$$

Мы видим, что в случае нормальных законов одно из наиболее важных свойств средних квадратических уклонений непосредственно переносится и на вероятные (срединные) уклонения.

§ 32. Решение задач

Условимся называть *основным нормальным законом* закон, для которого среднее значение равно нулю, а дисперсия — единице. Если x есть случайная величина, подчиненная основному нормальному закону, то условимся для краткости писать

$$\mathsf{P}\{|x| < a\} = \Phi(a)$$

для любого положительного числа a. Таким образом, $\Phi(a)$ есть вероятность того, что величина x, подчиненная основному нормальному закону, по абсолютному значению не превзойдет числа a. Для величины $\Phi(a)$ составлена очень точная таблица, дающая ее значения для различных значений числа a. Такая таблица служит незаменимым орудием для каждого, кому приходится иметь дело с расчетами вероятностей. Она прилагается ко всякой книге, посвященной теории вероятностей. В конце нашей книжки читатель также найдет такую таблицу. Имея таблицу значений функции $\Phi(a)$ под руками, можно легко и с большой точностью производить все расчеты для любых величин, распределенных по нормальным законам. Мы теперь покажем на примерах, как это делается.

Задача I. Случайная величина x распределена по нормальному закону со средним значением $\bar{x}$ и средним квадратическим уклонением Q_x. Найти вероятность того, что уклонение $x - \bar{x}$ по абсолютному значению не превзойдет числа a.

Пусть z — случайная величина, распределенная согласно основному нормальному закону. В силу основного свойства I (стр. 131) найдутся такие числа $c > 0$ и d, что величина $cz + d$ имеет среднее значение $\bar{x}$ и среднее квадратическое уклонение Q_x, т. е. подчинена тому же нормальному закону, что и данная величина x. Поэтому

$$\mathsf{P}(|x-\bar{x}|<a)=\mathsf{P}(|(cz+d)-(c\bar{z}+d)|<a)= \\ =\mathsf{P}(c|z-\bar{z}|<a);$$

но в силу формулы (2) на стр. 132 здесь $c=\frac{Q_x}{Q_z}=Q_x$, так как $Q_z = 1$ (для основного нормального закона дисперсии равна 1). Таким образом

$$\mathsf{P}(|x-\bar{x}|<a)=\mathsf{P}(Q_x|z-\bar{z}|<a)= \\ =\mathsf{P}\left(|z|<\frac{a}{Q_x}\right)=\Phi\left(\frac{a}{Q_x}\right). \qquad (4)$$

Этим поставленная задача решена, так как величина $\Phi\left(\frac{a}{Q_x}\right)$ непосредственно находится из таблицы. Таким образом, наша таблица с помощью формулы (4) позволяет легко вычислить вероятность любой границы уклонения для величины, подчиненной любому нормальному закону.

Пример 1. На станке изготовляется некоторая деталь. Оказывается, что ее длина x представляет собой случайную величину, распределенную по нормальному закону, и имеет среднее значение 20 *см* и дисперсию, равную 0,2 *см*. Найти вероятность того, что длина детали будет заключена между 19,7 *см* и 20,3 *см*, т. е. что уклонение в ту или другую сторону не превзойдет 0,3 *см*.

В силу формулы (4) и нашей таблицы

$$\mathsf{P}\{|x-20|<0{,}3\}=\Phi\left(\frac{0{,}3}{0{,}2}\right)=\Phi(1{,}5)=0{,}866.$$

Итак, около 87% всех изделий, изготовленных в данных условиях, будут иметь длины между 19,7 и 20,3 *см*; остальные 13% будут иметь бо́льшие отклонения от среднего.

Пример 2. В условиях примера 1 найти, какую точность длины изделия можно гарантировать с вероятностью 0,95.

Задача состоит, очевидно, в отыскании такого положительного числа a, для которого

$$\mathsf{P}\{|x-20|<a\}>0{,}95.$$

Расчет примера 1 показывает, что $a=0{,}3$ для этого мало, ибо в этом случае левая часть только что написанного неравенства меньше чем 0,87. Так как, согласно уравнению (4), $\mathsf{P}\{|x-20|<a\}=\Phi\left(\frac{a}{0{,}2}\right)=\Phi(5a)$, то нужно найти прежде всего в таблице такое значение $5a$, для которого $\Phi(5a)>0{,}95$. Находим, что это будет при $5a>1{,}97$, откуда $a>0{,}394$.

Таким образом, с вероятностью, превосходящей 0,95, можно гарантировать, что уклонение длины не превзойдет 0,4 *см*.

Пример 3. В некоторых практических вопросах считают, что случайная величина x, распределенная по нормальному закону, не обнаруживает уклонения большего, чем три средних квадратических уклонения Q_x. Какие основания имеются для этого утверждения?

Формула (4) и таблица показывают, что

$$\mathsf{P}\{|x-\bar{x}|<3Q_x\}=\Phi(3)>0{,}997,$$

и, следовательно, $\mathsf{P}\{|x-\bar{x}|>3Q_x\}<0{,}003$.

Практически это значит, что уклонения, превосходящие по абсолютному значению $3Q_x$, будут в среднем встречаться реже, чем три раза на тысячу. Можно ли пренебрегать такой возможностью или ее все же необходимо учитывать — это, конечно, зависит от содержания задачи и не может быть раз навсегда предписано.

Заметим, что соотношение $\mathsf{P}\{|x-\bar{x}|<3Q_x\}=\Phi(3)$ является, очевидно, частным случаем формулы

$$\mathsf{P}\{|x-\bar{x}|<aQ_x\}=\Phi(a), \tag{5}$$

вытекающей из формулы (4) и имеющей место для всякой случайной величины x, распределенной по нормальному закону.

Пример 4. При среднем весе некоторого изделия в 8,4 *кг* найдено, что уклонения, по абсолютному значению превосходящие 50 *г*, встречаются в среднем 3 раза на каждые 100 изделий. Допуская, что вес изделий распределен по нормальному закону, определить его вероятное уклонение.

Нам дано, что

$$\mathsf{P}(|x-8{,}4|>0{,}05)=0{,}03,$$

где x — вес случайно выбранного изделия. Отсюда

$$0{,}97=\mathsf{P}(|x-8{,}4|<0{,}05)=\Phi\left(\frac{0{,}05}{Q_x}\right);$$

таблица показывает, что $\Phi(a)=0{,}97$ при $a\approx 2{,}12$. Поэтому

$$\frac{0{,}05}{Q_x}\approx 2{,}12,$$

откуда

$$Q_x\approx\frac{0{,}05}{2{,}12}.$$

Вероятное отклонение, как мы знаем (стр. 134), равно

$$E_x=0{,}674Q_x\approx 0{,}0155\ \text{кг}=15{,}5\ \text{г}.$$

Пример 5. При стрельбе из орудия уклонение снаряда от цели вызывается тремя взаимно независимыми причинами: 1) погрешностью в определении положения цели, 2) погрешностью в наводке и 3) ошибкой от причин, меняющихся от выстрела к выстрелу (вес снаряда, атмосферные условия и т. п.). Предполагая, что все три погрешности распределены по нормальным законам со средним значением 0 и что вероятные уклонения их соответственно равны 24 *м*, 8 *м* и 12 *м*, найти вероятность того, что суммарное уклонение от цели не превзойдет 40 *м*.

Так как вероятное уклонение суммарной ошибки x в силу свойства III (стр. 135) равно

$$\sqrt{24^2+8^2+12^2}=28\ \text{м},$$

то среднее квадратическое уклонение суммарной ошибки равно

$$\frac{28}{0{,}674}\approx 41{,}5$$

и, значит,

$$\mathsf{P}(|x|<40)=\Phi\left(\frac{40}{41{,}5}\right)\approx\Phi(0{,}964)=0{,}665.$$

Уклонения, не превосходящие 40 *м*, будут, таким образом, наблюдаться примерно в $^2/_3$ всех случаев.

Задача II. Случайная величина x распределена по нормальному закону со средним значением $\bar{x}$ и средним квадратическим уклонением Q_x. Найти вероятность того, что уклонение $x-\bar{x}$ по абсолютному значению будет заключено между числами a и b $(0<a<b)$.

Так как по правилу сложения

$$\mathsf{P}(|x-\bar{x}|<b)=\mathsf{P}(|x-\bar{x}|<a)+\mathsf{P}(a<|x-\bar{x}|<b),$$

то

$$\mathsf{P}(a<|x-\bar{x}|<b)=\mathsf{P}(|x-\bar{x}|<b)-\mathsf{P}(|x-\bar{x}|<a)=\\=\Phi\left(\frac{b}{Q_x}\right)-\Phi\left(\frac{a}{Q_x}\right),\qquad(6)$$

чем и решается поставленная задача.

Для подавляющего большинства запросов практики та таблица значений величины $\Phi(a)$, которой мы все время пользовались, представляется, однако, излишне громоздким орудием расчета. Обычно бывает нужно рассчитать лишь вероятности попадания уклонения $x-\bar{x}$ в более или менее крупные отрезки; поэтому для целей практики желательно, наряду с нашей «полной» таблицей, иметь также и сокращенные таблицы, которые легко составить из полной таблицы с помощью формулы (6).

Приведем пример построения такого рода таблицы, гораздо более грубой, чем таблица в конце книги, и тем не менее во многих случаях совершенно достаточной. Разобьем весь промежуток изменения величины $|x-\bar{x}|$ на пять частей: 1) от нуля до $0,32Q_x$; 2) от

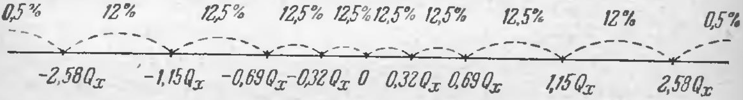

Рис. 16.

$0,32Q_x$ до $0,69Q_x$; 3) от $0,69Q_x$ до $1,15Q_x$; 4) от $1,15Q_x$ до $2,58Q_x$ и 5) свыше $2,58Q_x$.

Пользуясь формулой (4), находим:

$$\mathrm{P}(|x-\bar{x}|<0,32Q_x)=\Phi(0,32)\approx 0,25;$$

$$\mathrm{P}(0,32Q_x<|x-\bar{x}|<0,69Q_x)=\Phi(0,69)-\Phi(0,32)\approx 0,25;$$

$$\mathrm{P}(0,69Q_x<|x-\bar{x}|<1,15Q_x)=\Phi(1,15)-\Phi(0,69)\approx 0,25;$$

$$\mathrm{P}(1,15Q_x<|x-\bar{x}|<2,58Q_x)=\Phi(2,58)-\Phi(1,15)\approx 0,24;$$

$$\mathrm{P}(|x-\bar{x}|>2,58Q_x)=1-\Phi(2,58)\approx 0,01.$$

Результат этого расчета удобно изобразить с помощью графической схемы (рис. 16).

Здесь вся бесконечная прямая разделена на десять участков: пять положительных и пять отрицательных. Над каждым участком указано, какой процент фактически наблюдаемых уклонений будет в среднем падать на этот участок. Так, например, согласно вышеприведенному расчету на участки $(-1,15Q_x, -0,69Q_x)$ и $(0,69Q_x, 1,15Q_x)$ вместе взятые должно падать примерно 25% всех уклонений. В силу симметрии нормальных законов уклонения будут падать на оба участка приблизительно одинаково часто, так что на каждый из них будет ложиться около 12,5% общего числа уклонений. Имея под руками эту или подобную ей простую схему, мы можем непосредственно представить себе в основных чертах распределение уклонений для случайной величины, подчиненной нормальному закону с произвольными средним значением и средним квадратическим уклонением.

Рассмотрим, наконец, как вычисляется вероятность попадания случайной величины, подчиненной нормальному закону, в некоторый произвольно заданный отрезок.

Задача III. Зная, что случайная величина x распределена по нормальному закону (среднее значение $\bar{x}$, среднее квадратическое уклонение Q_x), вычислить с помощью таблицы вероятность неравенства $a < x < b$, где a и b $(a < b)$ — данные произвольные числа.

Нам придется рассмотреть три случая, в зависимости от расположения чисел a и b относительно $\bar{x}$.

Первый случай: $\bar{x} \leqslant a \leqslant b$.

По правилу сложения

$$\mathrm{P}(\bar{x} < x < b) = \mathrm{P}(\bar{x} < x < a) + \mathrm{P}(a < x < b),$$

откуда

$$\mathrm{P}(a < x < b) = \mathrm{P}(\bar{x} < x < b) - \mathrm{P}(\bar{x} < x < a) =$$
$$= \mathrm{P}(0 < x - \bar{x} < b - \bar{x}) - \mathrm{P}(0 < x - \bar{x} < a - \bar{x}).$$

Но при любом $\alpha > 0$, в силу симметрии нормальных законов:

$$\mathrm{P}(0 < x - \bar{x} < \alpha) = \mathrm{P}(-\alpha < x - \bar{x} < 0) =$$
$$= \frac{1}{2}\mathrm{P}(-\alpha < x - \bar{x} < \alpha) = \frac{1}{2}\mathrm{P}(|x - \bar{x}| < \alpha) =$$
$$= \frac{1}{2}\Phi\left(\frac{\alpha}{Q_x}\right); \qquad (7)$$

Поэтому

$$\mathsf{P}(a<x<b)=\frac{1}{2}\left\{\Phi\left(\frac{b-\bar{x}}{Q_x}\right)-\Phi\left(\frac{a-\bar{x}}{Q_x}\right)\right\}.$$

Второй случай: $a \leqslant \bar{x} \leqslant b$.

$$\begin{aligned}\mathsf{P}(a<x<b)&=\mathsf{P}(a<x<\bar{x})+\mathsf{P}(\bar{x}<x<b)=\\&=\mathsf{P}(a-\bar{x}<x-\bar{x}<0)+\mathsf{P}(0<x-\bar{x}<b-\bar{x})=\\&=\frac{1}{2}\left\{\Phi\left(\frac{\bar{x}-a}{Q_x}\right)+\Phi\left(\frac{b-\bar{x}}{Q_x}\right)\right\}\end{aligned}$$

в силу формулы (7).

Третий случай: $a \leqslant b \leqslant \bar{x}$.

$$\mathsf{P}(a<x<\bar{x})=\mathsf{P}(a<x<b)+\mathsf{P}(b<x<\bar{x}),$$

откуда

$$\begin{aligned}\mathsf{P}(a<x<b)&=\mathsf{P}(a<x<\bar{x})-\mathsf{P}(b<x<\bar{x})=\\&=\mathsf{P}(a-\bar{x}<x-\bar{x}<0)-\mathsf{P}(b-\bar{x}<x-\bar{x}<0)=\\&=\frac{1}{2}\left\{\Phi\left(\frac{\bar{x}-a}{Q_x}\right)-\Phi\left(\frac{\bar{x}-b}{Q_x}\right)\right\}.\end{aligned}$$

Задача решена во всех трех случаях. Мы видим, что для случайной величины, распределенной по любому нормальному закону, наша таблица дает возможность найти вероятность попадания этой величины в любой отрезок и тем самым исчерпывающим образом характеризует ее закон распределения. Для того чтобы посмотреть, как практически ведутся расчеты, рассмотрим следующий пример.

Пример. Стрельба ведется из точки O вдоль прямой OX. Средняя дальность полета снаряда равна 1200 *м*. Предполагая, что дальность полета H распределена по нормальному закону со средним квадратическим уклонением 40 *м*, найти, какой процент выпускаемых снарядов даст перелет от 60 до 80 *м*.

Для того чтобы снаряд имел такой перелет, мы должны иметь $1260 < H < 1280$; применяя заключительную формулу первого случая задачи III, мы находим:

$$P(1260 < H < 1280) = \\ = \frac{1}{2}\left\{\Phi\left(\frac{1280-1200}{40}\right) - \Phi\left(\frac{1260-1200}{40}\right)\right\} = \\ = \frac{1}{2}\{\Phi(2) - \Phi(1,5)\};$$

из таблицы находим:

$$\Phi(2) \approx 0{,}955, \qquad \Phi(1{,}5) \approx 0{,}866,$$

откуда

$$P(1260 < H < 1280) \approx 0{,}044;$$

мы видим, что с указанным перелетом будет падать немного более 4% выпускаемых снарядов.

ЧАСТЬ ТРЕТЬЯ
СЛУЧАЙНЫЕ ПРОЦЕССЫ

ГЛАВА ТРИНАДЦАТАЯ
ВВЕДЕНИЕ В ТЕОРИЮ СЛУЧАЙНЫХ ПРОЦЕССОВ

§ 33. Представление о случайном процессе

При изучении явлений природы, процессов техники, экономики или транспортных систем приходится часто встречаться с таким положением, что описание этих явлений или процессов производится с помощью случайных величин, которые изменяются во времени. Приведем несколько примеров.

Известно, что явление диффузии состоит в том, что молекулы одного вещества проникают в другое вещество и происходит перемешивание молекул разных веществ. Проследим за движением определенной молекулы. Пусть в начальный момент $t_0 = 0$ наблюдаемая нами молекула находилась в положении (x_0, y_0, z_0) и компоненты ее скорости по осям координат были (v_{0x}, v_{0y}, v_{0z}). В случайные моменты времени молекула сталкивается с другими молекулами и меняет не только свое положение, но также скорость и направление. Это изменение невозможно предсказать точно, поскольку нам не известны заранее ни моменты столкновений, ни их число за любой промежуток времени, ни скорости молекул, с которыми столкнется наша молекула. В результате положение данной молекулы в момент времени t определяется тремя координатами, $x(t)$, $y(t)$, $z(t)$, являющимися случайными функциями времени. Компоненты ее скорости $v_x(t)$, $v_y(t)$, $v_z(t)$ также представляют собой случайные величины, меняющиеся во времени.

Рассмотрим теперь сложное техническое устройство, состоящее из большого числа элементов — конденсаторов, сопротивлений, диодов, механических частей и т. д. Каждый элемент в силу тех или иных причин может терять свои рабочие свойства и приходить в состояние, в котором он перестает выполнять возложенные на него функции. Это состояние элемента мы назовем *отказом*. Длительные наблюдения над разнообразными техническими устройствами показывают, что время безотказной работы, т. е. время, прошедшее от момента начала работы до момента отказа, не может быть заранее точно указано, поскольку оно представляет собой случайную величину. Предположим теперь, что в момент отказа того или иного элемента отказавший элемент немедленно заменяется на такой же новый и исследуемое устройство продолжает свою работу. Спросим себя, сколько замен элементов приходится произвести за промежуток от 0 до t. Это число, которое обозначим через $n(t)$, зависит от t и является случайным. Мы получаем новый пример случайной величины, изменяющейся со временем. Эта случайная величина обладает одной особенностью: она не может убывать и изменяется в случайные моменты времени на целые числа (на число элементов, отказавших в момент замены). Такие случайные функции представляют значительный интерес в теории надежности — новой важной инженерной науке, широко использующей методы теории вероятностей.

В современной инженерной практике большую роль играет снабжение промышленных предприятий электрической энергией. Сколько энергии будет потреблено предприятием или цехом за данный срок? Как велика может быть потребляемая мощность в каждый данный момент? Как следует рассчитывать электрические кабели предприятий, чтобы они не были слишком маломощными и не перегорали от больших мощностей, которые нужно передавать в период нормальной работы? С другой стороны, как рассчитать потребляемую мощность, чтобы эти кабели не были чересчур большого сечения, поскольку при этом в кабели будет заложен излишний металл, а вместе с ним омертвлены значительные капиталы? Естественно, что для ответа на этот вопрос необходимо тщательно изучить реальную картину потребления электроэнергии как отдельными станками, механизмами, разного типа приборами и аппаратами, так и всеми потребителями тока, подключенными к одному кабелю (по терминологии энергетиков — фидеру). Такого типа исследования проведены на многих предприятиях: металлургических, металлообрабатывающих, нефтедобывающих, химических и иных. Мы приведем

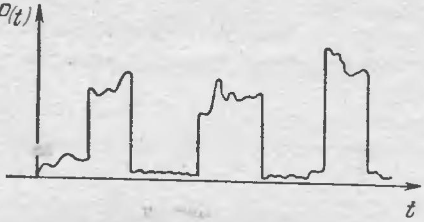

Рис. 17.

сейчас картину, типичную для металлообрабатывающих предприятий. Впрочем, суммарная картина будет такой же и для предприятий иного типа.

На рис. 17 изображена картина потребляемой мощности $P(t)$ токарным станком. Периоды, когда станок работает, сменяются периодами свободного хода, когда станок не несет полезной нагрузки. Потребляемая мощность в эти разные по характеру промежутки времени существенно различна: от почти нулевого потребления в свободных промежутках потребляемая мощность резко подскакивает в периоды работы. Однако при этом она не остается постоянной, а подвержена значительным колебаниям, поскольку в процессе обработки из-за локальной неоднородности обрабатываемого материала меняется скорость работы и режущие усилия. Одновременно оказывается, что совсем нерегулярно изменяются продолжительности рабочего и свободного периодов. При ближайшем и более тщательном изучении оказывается, что их изменение случайно. Мы вновь имеем дело со случайной функцией времени.

Если мы теперь запишем потребляемую мощность не одним станком, а большой группой 10—20 станков, то резкие колебания, которые имеют место на рис. 17, сгладятся. Суммарная потребляемая мощность не потеряет своего случайного характера, но приобретет большую плавность. В значительной степени это объясняется теми закономерностями, которые мы узнали при изучении закона больших чисел. Общий вид этой новой случайной функции дан на рис. 18. Процесс

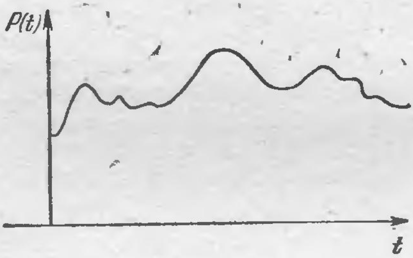

Рис. 18.

выравнивания связан с тем, что пиковые (максимальные) нагрузки одного приемника часто приходятся на те промежутки, когда другие приемники потребляют не слишком большую или даже минимальную мощность. Мы уже знаем, что дисперсия суммы n независимых случайных величин возрастает приблизительно как $\sqrt{n}$.

В настоящее время исследование электрических нагрузок промышленных предприятий, а также городских сетей все в большей мере базируется на указанных нами особенностях. Именно поэтому идеи, методы и математический аппарат теории вероятностей и теории случайных процессов (т. е. теории случайных функций одного независимого переменного) находят широкое применение при их решении.

§ 34. Понятие случайного процесса. Разные типы случайных процессов

Мы подошли к тому, чтобы дать определение случайного процесса. Представим себе, что некоторая случайная величина $\xi(t)$ зависит от непрерывно изменяющегося параметра t. Как правило, этот параметр называют временем, хотя в действительности он может иметь и другой смысл. Однако в подавляющем большинстве задач t действительно является временем.

Чтобы задать случайный процесс, нужно научиться не только описывать те значения, которые он может принимать в каждый момент времени, но и ожидаемые изменения принятых значений, а также вероятности возможных изменений процесса во времени и степень зависимости предстоящего развития процесса от его прошлой истории. Без этого говорить о том, что мы знаем случайный процесс, никак нельзя. Общий метод математического описания случайных процессов состоит в следующем: *для любого целого положительного числа n и для любых моментов времени $t_1, t_2, \ldots, t_n$ считаются известными функции*

$$F(t_1, t_2, \ldots, t_n, x_1, x_2, \ldots, x_n) = \\ = \mathrm{P}\{\xi(t_1) < x_1, \xi(t_2) < x_2, \ldots, \xi(t_n) < x_n\}.$$

Эти функции равны вероятностям одновременного выполнения неравенств $\xi(t_i) < x_i$ для всех выбранных моментов времени t_i, $i = 1, 2, \ldots, n$.

Предлагаемый метод описания случайных процессов универсален и позволяет в принципе выяснить все особенности поведения процесса во времени. Однако этот способ очень громоздок. Именно поэтому для получения более глубоких результатов приходится идти по иному пути: выделять важные частные типы случайных процессов и для них искать более совершенный аналитический аппарат, приспособленный к расчетам и построению математических моделей изучаемых явлений. В настоящее время, в связи с различными реальными процессами выделены несколько классов случайных процессов и их изучение продвинуто достаточно далеко. Необходимые для этого математические сведения выходят за пределы элементарных математических знаний.

Среди различных типов классов случайных процессов особое значение приобрели *марковские про-*

цессы, названные так в честь выдающегося русского математика конца XIX — начала XX столетия А. А. Маркова. Марков ввел в рассмотрение и первый стал систематически изучать свойства так называемых *цепных* зависимостей, явившихся прототипом для построения понятия и теории марковских случайных процессов.

Предположим, что процесс $\xi(t)$ обладает следующим свойством. Для любых моментов времени t_0 и t, $t_0 < t$, вероятность перейти из состояния x_0 в момент времени t_0 в состояние x (или в одно из состояний, принадлежащих некоторому множеству A) в момент t зависит только от t_0, x_0, t и x (или A), и дополнительное знание состояний, в которых был процесс в предшествующие t_0 моменты времени, ее не изменяет. Для таких процессов вся история их развития как бы концентрируется в достигнутом в момент t_0 состоянии x_0 и только через x_0 влияет на последующее его развитие. Именно такие процессы и называются *марковскими*.

Казалось бы на первый взгляд, что такая сильная схематизация явлений, которая заложена в марковских процессах, далека от действительных потребностей, поскольку обычно последействие предшествующей истории развития продолжается довольно долго. Однако опыт, накопленный математикой и ее применениями в биологии, технике, физике и других областях знания, убедительно показывает, что в схему марковских процессов прекрасно укладываются многие явления, такие, как явление диффузии или управление автоматизированным производством. Более того, оказалось, что путем изменения понятия состояния можно любой случайный процесс превратить в марковский. А это обстоятельство представляет очень серьезный аргумент в пользу широкого развития теории марковских процессов. Последнее замечание широко используется в исследовании многих вопросов практики, поскольку для марковских процессов удается воспользоваться хорошо разработанными сравнительно простыми аналитическими средствами расчета.

Добавим к сказанному, что любое использование математических средств исследования для изучения тех или иных явлений природы, технических, экономических или психических процессов требует предварительной их схематизации, выделения некоторых характерных особенностей, достаточно полно описывающих их течение. Теперь принято говорить не о схематизации явлений, а об их моделировании. Созданная нами модель явления обладает многими особенностями, облегчающими ее изучение. Во-первых, она проще самого изучаемого явления, во-вторых, для нее четко сформулированы исходные положения и связи, чего в реальных процессах, а особенно в явлениях экономики и биологии, не бывает. Изучив явление на достаточно простой модели и сравнив полученные выводы с результатами наблюдений самого явления, мы можем сделать заключение о качестве нашей модели и в случае необходимости внести в нее необходимые уточнения. В построении каждой математической модели неявно предполагается, что математический анализ применим к исследованию процесса изменения некоторой системы только в том случае, если каждое возможное состояние и эволюция исчерпывающе описываются посредством некоторого избранного нами математического аппарата.

Пожалуй, одной из самых замечательных математических моделей окружающих нас явлений определенной природы следует считать механику Ньютона. Простая схема протекания процессов и связанный с ней математический аппарат классических дифференциального и интегрального исчислений прекрасно описывают многочисленные процессы вот уже в течение четверти тысячелетия. Успехи машиностроения и первые полеты космических станций не только вблизи от Земли, но и к другим небесным телам, в значительной степени основаны на широком использовании классической механики Ньютона. В ней предполагается, что движение системы материальных точек полностью описывается положением и скоростью каждой из них. Иными словами, указание этих данных в момент t дает возможность однозначно вычислить состояние нашей системы для любого другого момента времени. Для этой цели механика предлагает уравнения движения.

Обратим внимание на то, что если бы мы под состоянием системы точек понимали только их положение в момент t, то такое понимание состояния было бы недостаточно для однозначного определения последующих состояний системы. Для механики Ньютона понятие состояния должно быть расширено добавлением значения скорости в данный момент.

Вне классической механики, собственно во всей современной физике приходится иметь дело со значительно более сложным положением дел, когда знание состояния системы в данный момент уже не может однозначно определить само состояние системы. Для марковских процессов однозначно определяется лишь вероятность перехода в то или иное состояние за заданный промежуток времени. Если угодно, то марковские процессы мы можем рассматривать как широкое обобщение процессов, изучаемых классической механикой.

§ 35. Простейший поток событий

Во многих практически важных или же интересных в познавательном отношении ситуациях приходится выяснять закономерности появления определенного типа событий: прибытие судов в морской порт, отказы в работе сложного устройства, замена перегоревших электрических лампочек, обрывы нитей на ватерной машине, регистрация моментов распада атомов радиоактивного вещества и т. д. Расчет многих предприятий бытового обслуживания — парикмахерских, касс магазинов, количества общественного транспорта, необходимого числа коек в больницах, пропускной способности шлюзов, переездов, мостов и т. д. тесно связан с изучением такого рода потоков. В последние годы было проведено тщательное изучение моментов прибытия самолетов в крупные аэропорты, прибытие грузовых судов в порты назначения, поступление вызовов на станции скорой медицинской помощи, поступление вызовов от абонентов телефонной сети на станцию и т. д. В результате этих наблюдений оказалось, что во всех указанных случаях с достаточно хорошим приближением появление этих событий хорошо описывается такой закономерностью.

Обозначим через t промежуток времени, который нас интересует, и положим, что $P_k(t)$ есть вероятность появления k событий потока за этот промежуток времени. Тогда при $k = 0, 1, 2, \ldots$ с большой точностью выполняется равенство

$$P_k(t) = \frac{(\lambda t)^n}{k!} e^{-\lambda t}, \qquad (1)$$

где λ — положительная постоянная, характеризующая собой «интенсивность» поступления событий потока. В частности, вероятность того, что за промежуток времени t не поступит ни одного события потока, равна

$$P_0(t) = e^{-\lambda t}. \qquad (2)$$

В молекулярной физике рассматривается задача: определить вероятность того, что в течение данного промежутка времени длительности t данная молекула не столкнется с другими молекулами. В книгах, посвященных соответствующим задачам физики, показывается, что эта вероятность как раз равна $e^{-\lambda t}$. Заметим, что если под потоком событий понимать в этом примере моменты столкновений данной молекулы с другими молекулами, то мы определяем как раз вероятность того, что в промежутке длины t не произойдет ни одного события потока.

Естественно предположить, что имеется какая-то общая причина, которая приводит к появлению одной и той же закономерности в столь различных по своему характеру явлениях. И действительно, оказалось, что имеются разнообразные и глубокие основания, в силу которых при весьма широких условиях только что описанная закономерность должна иметь место. Первая группа условий была найдена еще в начале нашего века известными физиками А. Эйнштейном и М. Смолуховским в связи с изучением броуновского движения.

Предположим, что изучаемый нами поток обладает следующими тремя свойствами:

1. Стационарностью. Под этим понимается следующее: *для любой группы конечного числа непересекающихся интервалов времени вероятность появления в них соответственно $k_1, k_2, \ldots, k_n$ событий зависит лишь от этих чисел и длин промежутков времени. В частности, вероятность появления k требований в промежутке $(T, t+T)$ не зависит от T и является функцией лишь k* и t.

2. Отсутствием последействия. Это свойство означает, что *вероятность поступления k событий потока в продолжение промежутка времени $(T, T+t)$ не зависит от того, сколько событий и как поступали до этого промежутка.* Это требование означает марковость изучаемого потока.

3. Ординарностью. Это условие выражает собой практическую *невозможность появления двух или большего числа событий за очень малый промежуток времени.*

Поток событий, удовлетворяющий этим трем условиям, называется *простейшим.*

Можно доказать, что простейший поток полностью определяется равенствами (1).

Простейший поток можно определить и иначе, а именно, как поток моментов времени, расстояния между которыми случайны, причем вероятность того, что расстояние между соседними моментами окажется большим, чем t, определяется формулой (2). Такое определение простейшего потока также часто используется при решении многих вопросов прикладного и теоретического характера.

Непосредственная проверка наличия трех перечисленных условий — стационарности, отсутствия последействия и ординарности — нередко трудно выполнима, поэтому очень важно найти иные условия, которые позволяли бы из иных оснований делать вывод о том, что поток событий окажется простейшим или близким к простейшему. Такое условие было найдено в работах ряда исследователей. Оно состоит в следующем.

Предположим, что интересующий нас поток является суммой очень большого числа независимых между собой стационарных потоков, каждый из которых лишь мало влияет на сумму. *Суммарный поток* при дополнительном ограничении арифметического характера, гарантирующего ординарность суммарного потока, *оказывается близким к простейшему.*

Эта теорема, доказанная в общей форме одним из создателей современной теории вероятностей А. Я. Хинчиным, представляет принципиальное значение для приложений.

Действительно, очень часто эта теорема дает возможность из общей структуры интересующего нас потока делать серьезные выводы. Так, из того, что поток вызовов, поступающих на телефонную станцию, можно рассматривать как сумму большого числа независимых потоков, каждый из которых ничтожно мало влияет на сумму, этот поток должен быть близок к простейшему. Точно так же поток грузовых судов, прибывающих в данный морской порт, составлен из большого числа потоков, отправляемых из различных других портов, следовательно, поток судов также должен быть близок к простейшему. Так оно и оказывается в действительности. Число примеров, имеющих иную реальную окраску, можно продолжать и далее.

§ 36. Одна задача теории массового обслуживания

Задача, которую мы здесь рассмотрим, является типичной для очень многих практически важных ситуаций. Опишем ее сначала в чисто прикладном плане, в каком она часто возникает перед работниками заводов, магазинов, складов, проектировщиками телефонных сетей.

Для удовлетворения некоторых потребностей населения организовано соответствующее предприятие — парикмахерская, телефонная станция, больница, зубоврачебная амбулатория и т. д. Требования на обслуживание поступают в случайные моменты времени и длительность их обслуживания также случайна. Спрашивается, как будут удовлетворены потребности клиентов, если оборудованы n мест обслуживания?

Условия, которые мы только что высказали, как легко видеть, хорошо отражают практическую ситуацию. Действительно, нет возможности указать, в какие моменты прибудут клиенты в парикмахерскую или в зубоврачебную амбулаторию. И нам хорошо известно, что нередко приходится ожидать очереди для получения необходимого нам обслуживания, а иногда его удается получить без всякого ожидания. Точно так же для выполнения, казалось бы, одной и той же операции требуется существенно различное время. При лечении зуба в зависимости от его состояния врач ограничивается лишь его очисткой или же сменой лекарства, а иногда проводит полностью за один прием всю процедуру пломбирования.

Естественно, что как клиентов, так и руководителей предприятия интересует в первую очередь такие характеристики качества обслуживания, как длина очереди на получение обслуживания, средняя длительность ожидания клиентом начала обслуживания, загруженность обслуживающих устройств, если нам известен средний темп поступления заявок на обслуживание и средний темп обслуживания.

При решении возникшей перед нами задачи будем исходить из следующих предположений: 1) поток требований на обслуживание является простейшим; 2) длительность обслуживания случайна и вероятность того, что на обслуживание придется затратить время, не меньшее чем t, равна $e^{-\nu t}$, где $\nu > 0$ — постоянная; 3) каждое требование обслуживается одним прибором; каждый прибор обслуживает только одно требование в момент, когда он занят; 4) если имеется очередь на обслуживание, то освободившийся прибор без потери времени переходит к обслуживанию очередного требования очереди.

Обозначим через $P_k(t)$ вероятность того, что в момент t в очереди находится k требований. В сформулированных нами условиях эти вероятности могут быть найдены при любом k, $k = 0, 1, 2, \ldots$ Однако точные формулы очень громоздки и на практике предпочитают пользоваться не ими, а теми, которые получаются из них для установившегося режима работы. Эти последние формулы несравненно проще и имеют следующий вид:

при $1 \leqslant k \leqslant n$

$$p_k = \frac{\rho^k}{k!} p_0; \tag{3}$$

при $k \geqslant n$

$$p_k = \frac{\rho^k}{n!\, n^{n-k}} p_0, \tag{4}$$

где

$$p_0 = \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!\,(n-\rho)}\right]^{-1} \quad \text{для} \quad \rho < n, \tag{5}$$

$$p_0 = 0 \quad \text{для} \quad \rho \geqslant n.$$

В этих формулах мы положили $\rho = \frac{\lambda}{\nu}$.

Обратим внимание на то, что при $\rho \geqslant n$ вероятность $p_0 = 0$. В силу формул (3) и (4) оказывается, что и при любом $k \geqslant 1$

$$p_k = 0.$$

Иными словами, при $\rho \geqslant n$ в установившемся процессе обслуживания застать в системе любое конечное число требований мы можем лишь с вероятностью нуль, т. е. с вероятностью единица в такой системе будет бесконечно много требований и образуется бесконечная очередь. Это означает следующее: во всех случаях, когда $\rho \geqslant n$, очередь на обслуживание неограниченно возрастает со временем.

Наш вывод имеет очень большое практическое значение, поскольку нередко при расчете количества необходимых средств обслуживания (посадочных площадок в аэропорту, причалов в морском порту, коек в больнице, пунктов выдачи инструмента на предприятии, касс в магазине и т. д.) принимают ложную предпосылку, что нужно исходить из «идеальной» производительности системы, равной отношению произведения числа обслуживающих приборов на длительность их использования в течение заданного срока к средней длительности одной операции обслуживания. Такой подсчет, как вытекает из сформулированного нами результата, в силу неравномерности поступления требований на обслуживание приводит к планированию очередей, а тем самым и к потерям времени, средств и потенциальных потребителей.

Конечно, умение использовать методы теории массового обслуживания приводит не только к тому, что мы получаем средство выяснять тот вред, который приносит перегрузка системы обслуживания, но и те потери, которые приносит излишнее количество обслуживающих средств в системе. Можно привести большое число примеров, когда теория массового обслуживания стала необходимой составной частью расчета крупных отраслей человеческой деятельности, например, телефонных станций, организации бригад ремонтных рабочих на предприятиях, работы крупных аэропортов или напряженных по движению автомобильных туннелей. Сейчас теория массового обслуживания приобретает все большее значение в вопросах конструирования современных электронных вычислительных машин, информационных систем, в ядерной физике, биологии и пр.

§ 37. Об одной задаче теории надежности

В последние годы во всем мире большое внимание стали уделять новой научной дисциплине, получившей название *теории надежности*. Ее цель состоит в том, чтобы разработать общие правила, которых следует придерживаться при проектировании, изготовлении, приемке, транспортировке, хранении и эксплуатации изделий промышленности, чтобы обеспечить наибольшую эффективность их использования. Естественно, что теория надежности разрабатывает также методы расчета надежности сложных изделий и технических систем по известным характеристикам надежности составляющих элементов.

Практическая важность этих задач не вызывает сомнений, поскольку вся наша жизнь связана прямо и косвенно с использованием разного типа технических устройств и систем. Мы каждый день пользуемся автобусами и трамваем для поездки на работу, выключателями для освещения наших квартир, водой из водопроводного крана, которую гонят по трубам машины, стоящие далеко от нас. В больницах употребляются разного типа устройства, помогающие восстановлению жизнедеятельности организма больных. Для примера, после операции на почках теперь применяется особый аппарат — искусственные почки, которые выполняют функции, свойственные естественным почкам в течение того периода, который необходим для восстановления их свойств. Ежегодно самолеты перевозят многие миллиарды пассажиров во все концы земного шара. И во всех случаях мы крайне заинтересованы в том, чтобы используемые нами технические средства выполняли возложенные на них функции абсолютно исправно на протяжении всей операции. Нарушение этого требования может привести к непоправимым последствиям. Действительно, самолет, двигатели которого отказали во время полета, грозит гибелью находящимся в нем пассажирам и членам экипажа. Искусственные почки, переставшие выполнять свои функции, грозят больному, который хорошо перенес операцию, гибелью в послеоперационный период. Авария, вызванная недостаточной надежностью насосов, установленных на водопроводной станции, нарушает снабжение города водой. В результате жители теряют возможность приготовления пищи, не имеют воды для питья и умывания; больницы вынуждены прекращать операции и нарушать нормальные процедуры лечения; улицы наполняются пылью, так как их нечем поливать и т. д.

Казалось бы, что все эти задачи не имеют никакого отношения к теории вероятностей и должны решаться инженерами-проектировщиками, коллективами работников заводов и эксплуатационников. Однако это не так, и на математиков ложится большая доля задач, связанных с исследованием количественной стороны расчетов, выработки целесообразных планов испытаний качества готовых изделий и получения заключений из результатов проведенных испытаний, расчетом оптимальных сроков проведения профилактических осмотров и ремонтов и т. д. И при этом оказывается, что все основные характеристики изделия, которые играют роль в его работе, носят вероятностный характер. Для примера, длительность безотказной работы одинаковых изделий, изготовленных на одном и том же заводе, из одного и того же сырья и в одних и тех же условиях, обладает значительным разбросом. Мы сами можем судить об этом достаточно определенно, если вспомним, как резко колеблются сроки службы электрических лампочек от момента их включения до момента их отправки в мусорный ящик. Мы знаем, что иногда лампочка работает безупречно годами, а иногда уже через несколько дней ее приходится заменять, поскольку у нее перегорел волосок. Длительные наблюдения и многочисленные специальные эксперименты убедительно показали, что мы не можем указать точно, какой срок в состоянии проработать данное изделие, а способны лишь оценить ту вероятность, с какой он проработает время, не меньшее чем заданное число t. Таким образом, теория вероятностей уверенно входит во все вопросы теории надежности и является одним из основных методов решения ее задач.

Перейдем теперь к рассмотрению одной несложной и притом в ее расчетной части схематической задаче. Мы поступаем так, чтобы не осложнять изложения, но в то же время достаточно ясно охарактеризовать саму стоящую перед нами практическую задачу.

Хорошо известно, что в природе нет абсолютно надежных элементов и изделий. Каждый элемент, как бы ни были совершенны его свойства, со временем теряет их. Для повышения надежности изделий поэтому приходится идти разными путями — облегчением условий эксплуатации, поиском более совершенных материалов, а также новых конструкций или же схем соединения. Одним из наиболее распространенных путей повышения надежности является резервирование. Сущность резервирования состоит в том, что в систему вводятся избыточные элементы, узлы и даже целые агрегаты, которые включаются в работу по мере выхода из рабочего состояния основных элементов. Для того чтобы обеспечить бесперебойное движение железнодорожных составов, на узловых станциях держат резервные паровозы и электровозы. На крупных электростанциях обязательно имеются дополнительные генераторы электрического тока. На особо ответственных линиях электроснабжения устраиваются параллельные линии, которые в нормальной обстановке используются неполностью. Каждый автомобиль помимо четырех колес имеет еще одно запасное — резервное.

Предположим, что имеется n устройств, которые должны работать одновременно в течение времени t. Вероятность того, что какое-нибудь из них безотказно проработает этот срок, равна p (одна и та же для всех устройств, и устройства отказывают независимо одно от другого). Отказ хотя бы одного из устройств приводит к отказу всей системы (например, прокол одного колеса автомобиля приводит к отказу всего автомобиля). Вероятность того, что наша система проработает безотказно, на основании формул Бернулли равна p^n.

Как изменится вероятность безотказной работы системы, если в ней, помимо n основных устройств, имеются еще m резервных, находящихся в нагруженном состоянии (т. е. в том же режиме, в каком и основные). Отказом по-прежнему считается переход системы в такое состояние, когда в ней число работоспособных устройств оказывается меньшим n. В силу теоремы сложения вероятностей искомая вероятность равна

$$\sum_{i=0}^{m} C_{m+n}^{n+i} p^{n+i} (1-p)^{m-i}.$$

Проиллюстрируем полученный результат несложным числовым примером. Пусть $n = 4$, $m = 1$, $p = 0{,}9$. Нетрудно подсчитать, что вероятность безотказной работы нерезервированной системы равна 0,6561. Если же у нас имеется один резервный элемент, то эта вероятность становится равной 0,9185. Таким образом, один-единственный элемент в резерве увеличивает вероятность безотказной работы четырех таких же элементов почти в полтора раза. Если бы у нас было два резервных элемента, то вероятность безотказной работы системы повысилась бы до 0,9841. Вот почему единственный резервный генератор на электростанции позволяет почти полностью исключить возможность выхода ее из рабочего состояния.

Надежность систем увеличивается во много раз, если используется так называемое *резервирование с восстановлением*. Каждое отказавшее устройство немедленно направляется на ремонт и сразу же после восстановления вновь поступает в резерв. Таким путем удается увеличить надежность резервированных систем во много раз.

Мы рассмотрели лишь упрощенную задачу теории резервирования. Исследование тех же вопросов, но в реальных условиях, требует привлечения гораздо более сложных математических средств, в первую очередь теории случайных процессов. Сейчас уже решены многие важные задачи теории надежности, но еще большое число вопросов далеко от удовлетворительного и окончательного решения. Систематическая работа позволит разрешить эти задачи в несколько упрощенных условиях и откроет пути их исследования в более реальных предпосылках.

ЗАКЛЮЧЕНИЕ

Теория вероятностей в последние десятилетия превратилась в одну из самых быстро развивающихся математических наук. Новые теоретические результаты открывают новые возможности для естественно-научного и практического использования методов теории вероятностей. Более тонкое, более детальное изучение явлений природы, а также производственных, технических, экономических и иных процессов толкает в то же время теорию вероятностей на разыскание новых методов, новых закономерностей, которые порождаются случаем. Теория вероятностей является одной из тех математических наук, которые не отгораживаются от жизни и от запросов других наук, а идут в ногу с общим развитием естествознания и техники. Читатель не должен превратно понять сказанное и подумать, что теория вероятностей превратилась теперь только в подспорье, во вспомогательное средство для решения прикладных задач. Отнюдь нет: за последние четыре десятилетия теория вероятностей превратилась в стройную математическую науку со своими проблемами и методами исследования. В то же время оказалось, что наиболее важные и естественные проблемы теории вероятностей как математической науки служат решению актуальных проблем естествознания.

Возникновение теории вероятностей относится к середине семнадцатого столетия и связано с именами Ферма (1601—1665), Паскаля (1623—1662) и Гюйгенса (1625—1695). В работах этих ученых в зачаточном виде появились понятия вероятности случайного события и математического ожидания случайной величины. Отправным пунктом для их исследований явились задачи, связанные с азартными играми. Однако важность новых понятий для изучения природы была им ясна, и, например, Гюйгенс в сочинении «О расчетах в азартных играх» писал: «Читатель заметит, что имеет дело не только с игрой, но что здесь закладываются основы очень интересной и глубокой теории». Среди позднейших ученых, оказавших значительное влияние на развитие теории вероятностей, необходимо указать Якова Бернулли (1654—1705), с именем которого мы уже встречались в тексте нашей книги, А. Муавра (1667—1754), Т. Байеса (ум. 1763), П. Лапласа (1749—1827), К. Гаусса (1777—1855) и С. Пуассона (1781—1840).

Мощное развитие теории вероятностей тесно связано с традициями и достижениями русской науки. В то время как в прошлом веке на Западе теория вероятностей зашла в тупик, в России гениальный математик П. Л. Чебышев нашел новый путь ее развития — всестороннее изучение последовательностей независимых случайных величин. Сам Чебышев и его ученики А. А. Марков и А. М. Ляпунов на этом пути получили фундаментальные результаты (закон больших чисел, теорема Ляпунова).

Читатели уже знакомы с законом больших чисел, и наша ближайшая задача состоит теперь в том, чтобы дать представление о другом важнейшем предложении теории вероятностей, получившем название теоремы Ляпунова (или иначе — центральной предельной теоремы теории вероятностей).

Причина столь большого ее значения заключается в том, что значительное количество явлений, исход которых зависит от случая, в основных своих чертах протекает по следующей схеме: изучаемое явление подвержено воздействию огромного количества независимо действующих случайных причин, каждая из которых оказывает лишь ничтожно малое влияние на течение явления в целом. Действие каждой из них выражается случайными величинам $\xi_1, \xi_2, \ldots, \xi_n$, а суммарное влияние на явление всех этих причин — суммой $s_n = \xi_1 + \xi_2 + \ldots + \xi_n$. Так как учет влияния каждой из этих причин (иначе говоря, указание функций распределения величин ξ) и даже простое перечисление их практически невозможно, то ясно, насколько важно разработать методы, позволяющие изучить их суммарное действие независимо от природы каждого отдельного слагаемого. Обычные методы исследования для решения поставленной задачи бессильны, на смену им должны прийти другие методы, для которых большое число действующих на явление причин было бы не препятствием, а, наоборот, облегчало бы решение поставленной задачи. Эти методы должны компенсировать недостаточное знание каждой отдельной действующей причины большим их числом, их массовостью. Центральная предельная теорема, установленная трудами, главным образом, академиков П. Л. Чебышева (1821—1894), А. А. Маркова (1856—1922) и А. М. Ляпунова (1857—1918), утверждает, что если действующие причины $\xi_1, \xi_2, \ldots, \xi_n$ взаимно независимы, их число n очень велико и действие каждой из этих причин по сравнению с суммарным их действием невелико, то закон распределения суммы s_n лишь незначительно может отличаться от нормального закона распределения.

Приведем примеры явлений, протекающих по только что описанной схеме.

При стрельбе из орудия по цели неизбежны отклонения точки попадания снаряда от точки прицеливания. Это — хорошо известное явление рассеивания снарядов. Так как рассеивание является результатом воздействия огромного числа независимо действующих факторов (например, неправильности в обточке стакана снаряда, головки снаряда, колебания в плотности материала, из которого выточена головка снаряда, ничтожные колебания количества взрывчатого вещества в различных снарядах, малые, незаметные для глаза ошибки в наводке орудия, ничтожные изменения состояния атмосферы при различных стрельбах и многие другие), каждый из которых лишь ничтожно мало влияет на траекторию снаряда, то из теоремы Ляпунова следует, что оно должно подчиняться нормальному закону. Это обстоятельство учитывается теорией стрельбы и кладется в основу при выработке правил стрельбы.

Когда мы производим какое-нибудь наблюдение с целью измерить ту или иную физическую константу, то на результат нашего наблюдения неизбежно влияет огромное количество факторов, каждый из которых в отдельности невозможно учесть, но которые порождают ошибки в измерении. Сюда относятся ошибки в состоянии измерительного прибора, показания которого могут нечувствительно меняться под влиянием различных атмосферных, тепловых, механических или иных причин. Сюда относятся ошибки наблюдателя, вызываемые особенностями его зрения или слуха и также нечувствительно меняющиеся в зависимости от психического или физического состояния наблюдателя. Фактическая ошибка измерения, таким образом, является результирующей огромного количества ничтожных по величине, независимых между собой, так сказать, элементарных, зависящих от случая ошибок. В силу теоремы Ляпунова мы вновь можем ожидать, что ошибки наблюдений будут подчинены нормальному закону распределения.

Подобных примеров можно привести сколько угодно: положение и скорость молекулы газа, определяемые большим числом столкновений с другими молекулами; количество продиффундировавшего вещества; уклонение деталей механизма от заданного размера при массовом производстве механизмов; распределение роста животных, растений или каких-либо из органов и т. д.

Совершенствование физической статистики, а также ряда отраслей техники поставило перед теорией вероятностей большое число совершенно новых проблем, не укладывающихся в рамки классических схем. В то время как физика или техника интересовало изучение **процесса**, т. е. явления, протекающего во времени, теория вероятностей не имела ни общих приемов, ни разработанных частных схем для решения возникающих при изучении таких явлений проблем. Возникла настоятельная потребность в разработке общей теории *случайных процессов*, т. е. теории, которая изучала бы случайные величины, *зависящие от одного или нескольких непрерывно изменяющихся параметров.*

Начало общей теории случайных процессов было положено фундаментальными работами советских математиков А. Н. Колмогорова и А. Я. Хинчина.

В известном смысле эта теория развивала введенные в первом десятилетии нашего века А. А. Марковым идеи изучения последовательностей зависимых случайных величин, получивших название цепей Маркова. Развитая им теория, вначале лишь как математическая дисциплина, в двадцатых годах в руках физиков превратилась в действенное орудие изучения природы. С тех пор многие ученые (С. Н. Бернштейн, В. И. Романовский, А. Н. Колмогоров, Г. Адамар, М. Фреше, В. Дёблин, Дж. Дуб, В. Феллер и др.) внесли в теорию цепей Маркова значительный вклад.

В двадцатых годах А. Н. Колмогоров, Е. Е. Слуцкий, А. Я. Хинчин и Поль Леви нашли тесную связь между теорией вероятностей и математическими дисциплинами, изучающими множества и общее понятие функции (теорией множеств и теорией функций действительного переменного). Несколько раньше к этим же идеям подошел Э. Борель. Открытие этой связи оказалось чрезвычайно плодотворным, и именно на этом пути удалось найти окончательное решение классических проблем, выдвинутых Чебышевым.

Наконец, следует отметить работы С. Н. Бернштейна, А. Н. Колмогорова и Мизеса по построению логически стройной теории вероятностей, способной к охвату разнообразных задач, возникающих перед ней в естествознании, технике и других областях знания.

В современном бурном развитии теории вероятностей особенно большую роль играет наука СССР, США, Франции, Великобритании, Швеции, Японии, Венгрии. Впрочем, интерес к этой научной дисциплине сильно возрос во всех странах в значительной мере под влиянием настойчивых требований практики в самых разнообразных ее проявлениях.

ПРИЛОЖЕНИЕ. ТАБЛИЦА ЗНАЧЕНИЙ ВЕЛИЧИНЫ Ф (*a*)

a	Ф (*a*)	*a*	Ф (*a*)	*a*	Ф (*a*)	*a*	Ф (*a*)	*a*	Ф (*a*)
0,00	0,000	0,60	0,451	1,20	0,770	1,80	0,928	2,40	0,984
0,01	0,008	0,61	0,458	1,21	0,774	1,81	0,930	2,41	0,984
0,02	0,016	0,62	0,465	1,22	0,778	1,82	0,931	2,42	0,984
0,03	0,024	0,63	0,471	1,23	0,781	1,83	0,933	2,43	0,985
0,04	0,032	0,64	0,478	1,24	0,785	1,84	0,934	2,44	0,985
0,05	0,040	0,65	0,484	1,25	0,789	1,85	0,936	2,45	0,986
0,06	0,048	0,66	0,491	1,26	0,792	1,86	0,937	2,46	0,986
0,07	0,056	0,67	0,497	1,27	0,796	1,87	0,939	2,47	0,986
0,08	0,064	0,68	0,504	1,28	0,800	1,88	0,940	2,48	0,987
0,09	0,072	0,69	0,510	1,29	0,803	1,89	0,941	2,49	0,987
0,10	0,080	0,70	0,516	1,30	0,806	1,90	0,943	2,50	0,988
0,11	0,088	0,71	0,522	1,31	0,810	1,91	0,944	2,51	0,988
0,12	0,096	0,72	0,528	1,32	0,813	1,92	0,945	2,52	0,988
0,13	0,103	0,73	0,535	1,33	0,816	1,93	0,946	2,53	0,989
0,14	0,111	0,74	0,541	1,34	0,820	1,94	0,948	2,54	0,989
0,15	0,119	0,75	0,547	1,35	0,823	1,95	0,949	2,55	0,989
0,16	0,127	0,76	0,553	1,36	0,826	1,96	0,950	2,56	0,990
0,17	0,135	0,77	0,559	1,37	0,829	1,97	0,951	2,57	0,990
0,18	0,143	0,78	0,565	1,38	0,832	1,98	0,952	2,58	0,990
0,19	0,151	0,79	0,570	1,39	0,835	1,99	0,953	2,59	0,990
0,20	0,159	0,80	0,676	1,40	0,838	2,00	0,955	2,60	0,991
0,21	0,166	0,81	0,582	1,41	0,841	2,01	0,956	2,61	0,991
0,22	0,174	0,82	0,588	1,42	0,844	2,02	0,957	2,62	0,991
0,23	0,182	0,83	0,593	1,43	0,847	2,03	0,958	2,63	0,991
0,24	0,190	0,84	0,599	1,44	0,850	2,04	0,959	2,64	0,992
0,25	0,197	0,85	0,605	1,45	0,853	2,05	0,960	2,65	0,992
0,26	0,205	0,86	0,610	1,46	0,856	2,06	0,961	2,66	0,992
0,27	0,213	0,87	0,616	1,47	0,858	2,07	0,962	2,67	0,992
0,28	0,221	0,88	0,621	1,48	0,861	2,08	0,962	2,68	0,993
0,29	0,228	0,89	0,627	1,49	0,864	2,09	0,963	2,69	0,993
0,30	0,236	0,90	0,632	1,50	0,866	2,10	0,964	2,70	0,993
0,31	0,243	0,91	0,637	1,51	0,867	2,11	0,965	2,72	0,993
0,32	0,251	0,92	0,642	1,52	0,871	2,12	0,966	2,74	0,994
0,33	0,259	0,93	0,648	1,53	0,874	2,13	0,967	2,76	0,994
0,34	0,266	0,94	0,653	1,54	0,876	2,14	0,968	2,78	0,995
0,35	0,274	0,95	0,658	1,55	0,879	2,15	0,968	2,80	0,995
0,36	0,281	0,96	0,663	1,56	0,881	2,16	0,969	2,82	0,995
0,37	0,289	0,97	0,668	1,57	0,884	2,17	0,970	2,84	0,995
0,38	0,296	0,98	0,673	1,58	0,886	2,18	0,971	2,86	0,996
0,39	0,303	0,99	0,678	1,69	0,888	2,19	0,971	2,88	0,996
0,40	0,311	1,00	0,683	1,60	0,890	2,20	0,972	2,90	0,996
0,41	0,318	1,01	0,688	1,61	0,893	2,21	0,973	2,92	0,996
0,42	0,326	1,02	0,692	1,62	0,895	2,22	0,974	2,94	0,997
0,43	0,333	1,03	0,697	1,63	0,897	2,23	0,974	2,96	0,997
0,44	0,340	1,04	0,702	1,64	0,899	2,24	0,975	2,98	0,997
0,45	0,347	1,05	0,706	1,65	0,901	2,25	0,976	3,00	0,997
0,46	0,354	1,06	0,711	1,66	0,903	2,26	0,976	3,10	0,998
0,47	0,362	1,07	0,715	1,67	0,905	2,27	0,977	3,20	0,999
0,48	0,369	1,08	0,720	1,68	0,907	2,28	0,977	3,30	0,999
0,49	0,376	1,09	0,724	1,69	0,909	2,29	0,978	3,40	0,999
0,50	0,383	1,10	0,729	1,70	0,911	2,30	0,979	3,50	0,9995
0,51	0,390	1,11	0,733	1,71	0,913	2,31	0,979	3,60	0,9997
0,52	0,397	1,12	0,737	1,72	0,915	2,32	0,980	3,70	0,9998
0,53	0,404	1,13	0,742	1,73	0,916	2,33	0,980	3,80	0,99986
0,54	0,411	1,14	0,746	1,74	0,918	2,34	0,981	3,90	0,99990
0,55	0,418	1,15	0,750	1,75	0,920	2,35	0,981		
0,56	0,425	1,16	0,754	1,76	0,922	2,36	0,982	4,00	0,99994
0,57	0,431	1,17	0,758	1,77	0,923	2,37	0,982		
0,58	0,438	1,18	0,762	1,78	0,925	2,38	0,983	5,00	0,99999994
0,59	0,445	1,19	0,766	1,79	0,927	2,39	0,983		

Борис Владимирович Гнеденко,
Александр Яковлевич Хинчин

Элементарное введение в теорию вероятностей

М., 1970 г., 168 стр. с илл.

Редактор *И. Е. Морозова*
Техн. редактор *В. С. Никифорова*
Корректор *Т. А. Панькова*

Сдано в набор 26/II 1970 г.
Подписано к печати 30/X 1970 г.
Бумага 84×108¹/₃₂. Физ. печ. л. 5,25.
Условн. печ. л. 8,82. Уч.-изд. л. 7,64. Тираж 100 000 экз.
Т-15469. Цена книги 23 коп. Заказ № 542.

Издательство «Наука»
Главная редакция
физико-математической литературы
Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени
Ленинградская типография № 2
имени Евгении Соколовой Главполиграфпрома
Комитета по печати при Совете Министров СССР.
Измайловский проспект, 29.