ЧАСТЬ ПЕРВАЯ
ОСНОВЫ МАТЕМАТИЧЕСКОГО АППАРАТА

ГЛАВА 1
ДИСКРЕТНОЕ ПРОСТРАНСТВО ЭЛЕМЕНТАРНЫХ СОБЫТИЙ

§ 1. Основные понятия

Мы будем заниматься теорией вероятностей в основном как одной из математических наук. Наука вообще, а математическая наука в особенности есть вещь сравнительно ясная, которая строится и излагается так, чтобы ее вполне сложившиеся части нетрудно было полностью понять. Но причины, по которым та или иная наука существует, преподается и развивается, коренятся в малодоступных анализу глубинах индивидуальной и общественной психологии. Из многих источников хорошо известно, что с давних пор внимание человека привлекали «эксперименты» (как мы говорим сейчас), исход которых не вполне однозначен: эксперимент может кончиться одним из исходов, полный список которых обозначим $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Например *, в греческой и римской цивилизациях важным элементом общественной жизни было гадание (в частности, по внутренностям жертвенных животных); впрочем, сомнительно, чтобы существовал полный список возможных исходов такого «эксперимента». Различные варианты средневекового «божьего суда», например судебный поединок, а в наше время жеребьевка для решения тех или иных спорных вопросов; наконец, азартные игры всех видов — все

---

* В данной книге мы неоднократно будем обращаться к тем или иным историческим сведениям, в особенности, к истории теории вероятностей. Однако систематического характера эти сведения не носят. Причина их фрагментарности состоит главным образом в том, что на общепринятые исторические взгляды нельзя полагаться без их фундаментальной проверки. Ярким примером такого рода является, например, знаменитая формула Лейбница—Ньютона, которая, как оказывается, была настолько хорошо известна до Лейбница и Ньютона, что эти два великих ученых никогда из-за нее и не спорили, хотя, как кажется, не упустили ни одной возможности для приоритетных споров, нелепый характер которых служит вечным назиданием потомству. Автор данной книги не занимался необходимой переоценкой исторических взглядов, следовательно, и не мог браться за систематическое изложение истории теории вероятностей. Читателю, заинтересованному в подобном систематическом изложении, горячо рекомендуется обратиться к историческому очерку в учебнике Б. В. Гнеденко (шестое издание 1988 г.), который содержит много интересного.

это примеры «экспериментов» с неопределенным исходом. Например, при бросании монеты она может упасть вверх гербом (что мы будем изображать единицей) или цифрой (что мы будем изображать нулем). Не следует непочтительно относиться к бросанию монеты: из многих бросаний можно сконструировать достаточно интересные эксперименты, а из счетного числа — даже бросание случайной точки на континуум [0, 1]. Для этого нужно на нули и единицы, возникающие при бросании монеты, посмотреть как на знаки двоичной дроби, определяющей вещественное число.

Понятия теории вероятностей применимы не ко всем экспериментам с неопределенным исходом. Давно было замечено, что эксперименты, производимые с помощью достаточно аккуратно сделанных «аппаратов», как, например, монета, игральная кость, рулетка или колода игральных карт, обладают двумя свойствами: 1) непредсказуемостью (в смысле невозможности заранее предсказать исход такого эксперимента); 2) статистической устойчивостью: при большом числе повторений эксперимента частота осуществления того или иного исхода оказывается близкой к некоторому числу, которое и называют вероятностью данного исхода. (Частотой называется отношение числа наступлений данного исхода к числу всех экспериментов).

Иногда вероятности исходов можно угадать из соображений симметрии. Так, для монеты вероятность выпадения герба, очевидно, должна быть такой же, как и для выпадения цифры, т. е. равняться $1/2$. Опыты с реальными монетами это подтверждают. Аналогично для игральной кости (кубик с шестью гранями, на которых нанесены точки числом от 1 до 6) вероятность выпадения каждой грани должна равняться $1/6$. Но опыты это не всегда подтверждают. Впрочем, эти (вообще говоря, не равные между собой) вероятности все-таки оказываются близкими к $1/6$, так что в учебных задачах с хорошим приближением считается, что мы имеем дело с идеальной костью с вероятностями выпадения отдельных граней, равными $1/6$.

Анализ классических монографий и учебников по теории вероятностей (начиная от Лапласа, Пуассона, Чебышева и до наших дней) показывает, что все они начинаются с одной и той же математической модели случайного эксперимента, в которой считается заданным множество Ω элементарных исходов эксперимента и вероятности $\mathbf{P}(\omega)$ каждого элементарного исхода $\omega\in\Omega$. В классических учебниках Ω конечно, а в современных (по ряду разумных причин) счетно.

Определение 1. Вероятностным пространством называется не более чем счетное множество $\Omega=\{\omega\}$, каждому элементу ω которого поставлено в соответствие число $\mathbf{P}(\omega)\geqslant 0$, называемое вероятностью ω.

При этом должна выполняться единственная

Аксиома.

$$\sum_{\omega\in\Omega} \mathbf{P}(\omega) = 1$$

(сумма вероятностей всех элементарных исходов равна единице).

Кроме элементарных исходов эксперимента (синоним: *элементарные события*) в теории вероятностей выделяются *события,* которые в классических учебниках задавались словесным описанием, например событие, состоящее в том, что при бросании кости выпадает четное число очков. Постепенно было осознано, что событие лучше всего определить как произвольное подмножество множества Ω.

Определение 2. Событием A называется произвольное подмножество $A\subseteq\Omega$. Вероятность $\mathbf{P}(A)$ события A определяется формулой

$$\mathbf{P}(A) = \sum_{\omega\in A} \mathbf{P}(\omega)$$

(вероятность события есть сумма вероятностей входящих в него элементарных исходов).

Совокупность определений 1 и 2 и единственной аксиомы дает полное описание элементарной модели теории вероятностей. В принципе эта модель достаточна для решения всех задач, связанных со случаями, когда множество элементарных событий дискретно (т. е. конечно или счетно), но перед изучающим теорию вероятностей стоит (вообще говоря, нелегкая) задача перевода формулировок ситуаций, заданных в терминах обычного языка, на язык вероятностного пространства, т. е. Ω и $\mathbf{P}(\omega)$. Если взглянуть на проблему изучения теории вероятностей шире — изучение с целью применения к реальным явлениям, — то вырисовывается еще более сложная картина.

Дело в том, что любая математическая наука (теория вероятностей в том числе) не несет в самой себе никаких указаний на возможные области и способы применений. Например, при изучении математического анализа функций нескольких переменных, векторных (и тензорных) полей совершенно ниоткуда не следует, что этот аппарат находит применение в электродинамике. А скажем, уравнение струны в уравнениях математической физики изучают никак не ради скрипичной струны. Задачник по теории вероятностей суммирует (насколько сумеет) накопленный веками опыт применения этой науки к реальным явлениям. Поэтому ситуации задачника не могут и не должны формулироваться в чисто математических терминах. В реальной научной деятельности предполагается двухэтапный перевод: реальной ситуации в ситуацию задачника, затем ситуации задачника в ситуацию Ω, $\mathbf{P}(\omega)$. В учебном процессе изучается в основном лишь второй этап.

Надо отметить, что описание множества Ω обычно не представляет трудностей: это просто множество всех возможных исходов эксперимента. Но задача определения $\mathbf{P}(\omega)$ часто является трудной. Существует классический прием, так называемая «классическая вероятность», когда множество Ω конечно, а все $\mathbf{P}(\omega)$ равны между собой (в этом случае $\mathbf{P}(\omega) = 1/N$, где $N = N(\Omega)$ — число элементов в множестве Ω). Для любого события A в этом случае имеем

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\omega) = \frac{N(A)}{N(\Omega)},$$

где $N(A)$ — число элементов в множестве A.

На классическом языке все $\omega \in A$ называются элементарными исходами, *благоприятными* для события A, и получаем классическое «заклинание»: ***вероятность события равна отношению числа исходов, благоприятных для данного события, к числу всех возможных исходов.***

Следует отметить, что классики теории вероятностей, скажем Лаплас, прекрасно понимали, что элементарные исходы могут быть и не одинаково вероятными. Но «классическая вероятность» закрепилась в науке как прием, позволяющий быстро и легко (хотя, быть может, и неверно) решить задачу об определении $\mathbf{P}(\omega)$. Это решение обычно мотивируется теми или иными соображениями о «симметрии», т. е. соображениями теоретико-групповыми.

Рассмотрим *пример* древней процедуры жеребьевки, которая возобновляется и в наши дни при каждой сдаче экзамена группой студентов. Пусть имеется N экзаменационных билетов, из которых n «счастливых» (в том смысле, что все студенты их знают), а $N-n$ «несчастливых» (т. е. ни один из студентов их не знает), причем для простоты обозначений всех студентов тоже N. Жеребьевка (т. е. раздача билетов) происходит по очереди: сначала берет билет первый в очереди студент, затем второй и т. д. Понятно, что для первого студента вероятность вытащить счастливый билет равна n/N, но как быть со вторым? Если первый студент вытащит счастливый билет, то шансы второго составят $(n-1)/(N-1)$, т. е. уменьшатся, а если первый студент вытащит несчастливый билет, то шансы второго будут $n/(N-1) > n/N$. Для третьего студента нужно рассмотреть еще более сложный набор ситуаций и т. д. Пусть A_j — событие, состоящее в том, что j-й в очереди студент вытащит счастливый билет. Попытаемся найти $\mathbf{P}(A_j)$, введя множество Ω элементарных событий так, чтобы они были равновероятными.

Предлагается под отдельным элементарным событием ω понимать тот список, который окажется в руках у экзаменатора после окончания раздачи билетов:

$$\omega = \begin{pmatrix} 1 & 2 & \ldots & j & \ldots & N \\ i_1 & i_2 & \ldots & i_j & \ldots & i_N \end{pmatrix}$$

(в первой строке — номера студентов, во второй строке — номера билетов). Если угодно, это подстановка из N чисел. Номера билетов $i_1, i_2, \ldots, i_N$ как-то зависят от того порядка, в котором их разложил на столе экзаменатор. Можно предположить, что экзаменатор положил сверху счастливые билеты, а несчастливые засунул под них. Тогда первым студентам в очереди будет лучше (если они не будут хитрить, а возьмут попросту те билеты, которые лежат сверху). Можно предположить, что экзаменатор сделал наоборот. Оба этих случая не приведут к задаче на классическую вероятность. Но если предположить, что экзаменатор подобными пустяками не занимается, то тогда задача инвариантна относительно любых перестановок номеров билетов. Но перестановка номеров i_1 и i_j эквивалентна перестановке первого и j-го студентов, так что ясно, что должно быть $\mathbf{P}(A_j) = \mathbf{P}(A_1) = n/N$.

Можно сосчитать эти вероятности и из классической формулы. Очевидно, что $N(\Omega) = N!$ Для подсчета $N(A_j)$ заметим, что для $\omega \in A_j$ номер i_j может принимать n различных значений; i_1 — все значения, кроме i_j, т. е. $N-1$ значений, i_2 — все значения, кроме i_j и i_1, т. е. $N-2$ значений и т. д. Поэтому $N(A_j) = n(N-1)!$ и $\mathbf{P}(A_j) = N(A_j)/N(\Omega) = n/N$.

Таким образом, вероятность вытащить счастливый билет не зависит от места в очереди: не нужно ни приходить пораньше, чтобы занять очередь, ни стараться оказаться в конце очереди. Этот вывод целиком зависит от предполагаемой равновероятности элементарных событий. Каким образом можно было бы обосновать это допущение?

Имея в виду связь между вероятностью и частотой, можно было бы представить себе экспериментальную проверку: в длинном ряде экспериментов определяются частоты наступления различных элементарных событий; если эти частоты оказываются близкими, то мы заключаем, что соответствующие вероятности в самом деле равны. Однако элементарных событий у нас $N!$; чтобы частоты их наступления сделались похожи на вероятности, необходимо провести столько экспериментов, чтобы каждое элементарное событие произошло хотя бы несколько раз, т. е., скажем, $10N!$ или $100N!$ экспериментов. Ясно, что это совершенно невозможно уже при умеренном N. Следовательно, мысль об экспериментальной проверке равновероятности должна быть оставлена.

Теоретические соображения о равновероятности сводились у нас к тому, что экзаменатор не станет заниматься такими пустяками, как создание заведомо неравновероятной ситуации. Это совсем не означает, что автоматически создается равновероятность: например, опыты с игральными картами показали, что колоду карт нужно очень долго тасовать, чтобы достаточно хорошо разрушить какой-то первоначальный порядок. С экзаменационными билетами этого никто не делает.

Поэтому правильнее будет сказать, что если различные элементарные события неравновероятны, то ни экзаменатор, ни студенты совершенно не знают, какие из этих событий имеют большие вероятности, а какие — меньшие. Это незнание мы и моделируем в математической модели с равновероятными элементарными событиями. Если угодно, наша равновероятность в этой задаче субъективная; несмотря на все попытки изгнать субъективизм из науки, что-нибудь от него всегда остается.

§ 2. Исчисление вероятностей

В предыдущем параграфе был приведен пример практического вывода, который можно извлечь из подсчета вероятностей: очередь за экзаменационными билетами занимать смысла не имеет. Какие выводы можно вообще извлечь из знания вероятностей, пока сказать не можем: это станет ясным лишь после изучения основных законов теории вероятностей. Пока что наша задача — облегчить подсчет вероятностей. Дело в том, что прямой подсчет с помощью определения 2 § 1, конечно, часто бывает трудным. Следует вывести ряд простых формул, вытекающих из определения 2, которые будем называть формулами исчисления вероятностей.

Под исчислением вообще понимаются какие-то способы писать формулы и выводить из одних формул другие. Огромную нагрузку несут в современной культуре дифференциальное и интегральное исчисления. Роль исчисления вероятностей не столь фундаментальна, но все же велика. Кроме облегчения подсчетов вероятностей, можно отметить два аспекта работы того небольшого исчисления, которое сейчас разовьем.

1. Выведенные для дискретного Ω формулы исчисления сохраняют свой вид для произвольного сложного Ω, облегчая понимание общей аксиоматики.

2. Лишь в учебниках теории вероятностей действует схема, согласно которой для реального явления нужно сначала построить модель из Ω и $\mathbf{P}(\omega)$ и воспользоваться определением 2 § 1. Фактически вероятности одних событий находятся по вероятностям других событий без полного описания (иногда) Ω и (как правило) $\mathbf{P}(\omega)$, а прямо путем обращения к формулам исчисления.

Итак, рассмотрим *операции над событиями и свойства вероятностей*. Поскольку события трактуются как подмножества Ω, то операции над ними — это обычные теоретико-множественные операции, но следует иметь в виду, что в теории вероятностей сохранилась старая терминология, употреблявшаяся еще до возникновения теории множеств (в этом есть свой смысл, так как в применениях теории вероятностей событие есть, конечно, подмножество Ω, но, как правило, не любое, а задающееся достаточно простым высказыванием).

Дополнение (или *отрицание*, или *противоположное событие*). С каждым событием A связано событие $\bar{A}=\Omega\backslash A$, которое состоит из тех и только тех элементарных событий $\omega\in\Omega$, которые не входят в A. Это событие $\bar{A}$ называется *дополнением* к событию A, либо *отрицанием* события A, либо событием, *противоположным* для события A. Из определения 2 § 1 вытекает, что

$$\mathbf{P}(A)+\mathbf{P}(\bar{A})=1,$$

так как

$$\mathbf{P}(A)+\mathbf{P}(\bar{A})=\sum_{\omega\in A}\mathbf{P}(\omega)+\sum_{\omega\notin A}\mathbf{P}(\omega)=\sum_{\omega\in\Omega}\mathbf{P}(\omega)=1.$$

Поупражняемся немного в классической терминологии теории вероятностей. Будем говорить, что событие A наступило в опыте, если опыт закончился таким элементарным событием ω, что $\omega\in A$. Тогда сможем заявить, что противоположное событие $\bar{A}$ наступает тогда, когда событие A не наступает: во всяком опыте наступает A или $\bar{A}$, но никогда оба вместе.

Объединение, или сумма. Суммой событий $A\cup B$ называется теоретико-множественное объединение соответствующих подмножеств A и B.

Пересечение, или произведение. Произведением $A\cap B=AB$ называется теоретико-множественное пересечение подмножеств A и B.

Очевидно, что *сумма событий* наступает тогда, когда наступает *хотя бы одно из них*, а *произведение* — тогда, когда наступают *оба вместе*. Имеет место формула

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

Действительно, в сумме $\mathbf{P}(A)+\mathbf{P}(B)$ вероятности элементарных событий, входящих и в A и в B, будут сосчитаны дважды; но если теперь вычесть $\mathbf{P}(AB)$, то остается сумма вероятностей элементарных событий, входящих в A и B, в которой каждое элементарное событие сосчитано ровно один раз. А это и есть $\mathbf{P}(A\cup B)$.

В этой книге мы резервируем знак «плюс» для обозначения суммы непересекающихся множеств: будем писать вместо $A\cup B$ сумму $A+B$, если известно, что пересечение AB пусто: $AB=\varnothing$. В таком случае $\mathbf{P}(AB)=0$ и получаем

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

В элементарном случае (когда Ω не более чем счетно) эта формула является простенькой теоремой; она сохраняется и для общего случая, но уже в качестве аксиомы.

Свойств операций над событиями можно отметить великое множество, например дополнение к сумме событий равно пересечению их дополнений; дополнение же к пересечению есть, наоборот, сумма дополнений. Свойствам операций отвечают и какие-то свойства вероятностей. Однако основные формулы исчисления, которые исчерпывают значительную часть того, что обычно требуется в выкладках, уже приведены.

Рассмотрим теперь следующую *задачу*. Некто написал n писем, предназначенных n различным адресатам; затем на конвертах написал n адресов и случайно разложил письма по конвертам. Какова вероятность того, что *хотя бы одно* письмо попало в свой конверт?

Элементарные события здесь, очевидно, подстановки; слова «случайно разложил» обозначают, что эти подстановки равновероятны. Спрашивается, сколько таких подстановок, в которых хотя бы один символ переходит в себя. Этот подсчет может показаться затруднительным. Воспользуемся формулами исчисления. Пусть событие A_i означает, что i-е письмо попало в свой конверт: очевидно, что $\mathbf{P}(A_i)=1/n$. Нас спрашивают о вероятности суммы $A_1\cup A_2\cup\ldots\cup A_n$. Оказывается, что имеет место формула

$$\mathbf{P}(A_1\cup A_2\cup\ldots\cup A_n)=\sum_i \mathbf{P}(A_i)-\sum_{i<j}\mathbf{P}(A_iA_j)+$$
$$+\sum_{i<j<k}\mathbf{P}(A_iA_jA_k)-\ldots\pm\mathbf{P}(A_1A_2\ldots A_n).$$

Докажем эту формулу. Для этого нужно доказать, что в сумме, стоящей в правой части, вероятность $\mathbf{P}(\omega)$ каждого $\omega\in A_1\cup A_2\cup\ldots\cup A_n$ учтется ровно один раз. Пусть для определенности, $\omega\in A_1A_2\ldots A_k$, но $\omega\notin A_{k+1},\ldots,\omega\notin A_n$. Тогда в правую часть $\mathbf{P}(\omega)$ войдет следующее число раз:

$$C_k^1-C_k^2+C_k^3-\ldots\pm C_k^k=1,$$

так как

$$0=(1-1)^k=1-C_k^1+C_k^2-C_k^3+\ldots\pm C_k^k.$$

Формула доказана.

Заметим теперь, что

$$\mathbf{P}(A_1A_2\ldots A_k)=\frac{(n-k)!}{n!},$$

так как событие $A_1A_2\ldots A_k$ означает, что первые k писем попали в свои конверты, а остальные $n-k$ переставились как угодно, Учитывая это, находим, что

$$\mathbf{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = C_n^1 \frac{(n-1)!}{n!} - C_n^2 \frac{(n-2)!}{n!} + \ldots \pm C_n^n \frac{1}{n!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots \pm \frac{1}{n!},$$

в чем нетрудно узнать отрезок ряда для $1\text{-}e^{-1} \approx {}^2/_3$.

Таким образом, при большом n вероятность каждому отдельному письму попасть в свой конверт весьма мала, но из большего числа маловероятных событий хотя бы одно произойти вполне может. Этот вывод мы будем развивать неоднократно.

§ 3. Условная вероятность

Малоинтересно рассматривать одно случайное событие: у такого события закон один — может наступить или не наступить. Сколько-нибудь содержательная наука начинается тогда, когда в рассмотрение вводится много событий. В математической схеме все они являются подмножествами одного Ω (определение 1 § 1), но в реальной жизни нужно уметь выбрать Ω и определить $\mathbf{P}(\omega)$ для $\omega \in \Omega$. Если мы ввели в рассмотрение какие-то события $A_1, A_2, \ldots, A_n$, то должны иметь право рассматривать и их комбинации (т. е. то, что получится из них операциями дополнения, суммы и пересечения). Отсюда вытекает, что наименьшее Ω, пригодное для описания n событий A_1, $A_2, \ldots, A_n$, состоит из элементарных событий вида $B_1 B_2 \ldots, B_n$, где каждое B_i может принимать два значения A_i и $\bar{A}_i$. (Иными словами, если мы ввели в рассмотрение события $A_1, A_2, \ldots, A_n$, то мы ввели в рассмотрение все ситуации, когда некоторые из этих событий происходят, а некоторые — не происходят). Следовательно, нужно уметь определить вероятности $\mathbf{P}(B_1 B_2 \ldots B_n)$.

Как для этой цели, так и для ряда других применяется понятие условной вероятности. Математический термин часто не имеет ничего общего с общеязыковым смыслом соответствующего слова (вспомните, например, алгебраические «кольца», «поля», «идеалы»). Но в случае термина «условная вероятность» это, к счастью, не так — математический термин точно соответствует общеязыковому смыслу высказывания: «вероятность того, что событие B произойдет, если известно, что событие A произошло».

Вдумаемся, действительно, сначала в общеязыковый смысл слов «условная частота события B при условии, что событие A произошло». Нужно представить себе длинный ряд из n опытов, в котором событие A произошло n_A раз, а событие B — n_B раз. Тогда частота события A равна n_A/n, а частота события B равна n_B/n. Что же такое условная частота события B при условии, что A произошло? Очевидно, из всех n опытов нужно рассмотреть лишь те n_A опытов, в которых A произошло, и вычислить частоту наступления B в этих опытах. Но число наступлений события B в этих опытах есть, очевидно, число наступлений события AB во всех опытах, и его естественно обозначить n_{AB}. Итак, условная частота события B при условии, что A произошло, есть

$$n_{AB}/n_A = (n_{AB}/n) \ / \ (n_A/n).$$

Но если число n всех опытов достаточно велико, то частоты должны быть близки к вероятностям: $n_{AB}/n \approx \mathbf{P}(AB)$ и $n_A/n \approx \mathbf{P}(A)$. Примем поэтому *математическое определение* условной вероятности $\mathbf{P}(B/A)$ события B при условии, что событие A произошло, в следующем виде:

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

(«вероятность совместного наступления, деленная на вероятность условия»). Предполагается, конечно, что $\mathbf{P}(A) \neq 0$.

Учебники (особенно старые) из этого определения считают нужным вывести «теорему умножения вероятностей» в виде

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

сопровождая ее «заклинанием» следующего вида: «вероятность совместного наступления двух событий равна вероятности одного из них, умноженной на условную вероятность второго при условии, что первое наступило».

Не следует видеть в этом лишь пристрастие к многократному переписыванию тривиальностей. Это тривиальности, если мы каким-то образом ввели безусловные вероятности, но на практике введение безусловных вероятностей обычно трудно.

Разберем, например, знаменитую *ошибку* прекрасного ученого *Д'Аламбера*. Речь идет об одновременном бросании двух монет. В результате этого опыта либо: 1) обе монеты выпадают гербом вверх (ГГ); 2) обе монеты выпадают цифрой вверх (ЦЦ), наконец, 3) монеты выпадают разными сторонами кверху. Д'Аламбер считал, что эти результаты опыта равновероятны (следовательно, вероятность каждого равна $1/3$). Между тем и до Д'Аламбера, и после Д'Аламбера существовал правильный взгляд на вещи, который состоит в том, что мысленно следует сделать монеты (а также кости и т. п.) различимыми; тогда исход 3) представится как объединение исходов ГЦ$\cup$ЦГ; четыре исхода ГГ, ЦЦ, ГЦ, ЦГ нужно считать равновероятными; в частности, исход 3) имеет вероятность $1/2$.

Следовало бы обосновать, почему не прав Д'Аламбер. Насколько понимает автор данной книги, сделать это чисто логическим путем невозможно: придется в конце концов промямлить что-нибудь вроде того, что мнение Д'Аламбера не согласуется с опытом. Действительно, в физике микромира бывают ситуации, в которых скорее прав Д'Аламбер. Но надо хорошо представлять себе, к какому именно опыту мы апеллируем; нам ведь слишком скучно бросать монеты, чтобы речь могла идти о личном опыте. Речь идет о некотором исторически накопленном опыте игроков в орлянку или кости (вроде знаменитого шевалье де Мере), которым почти бессознательно пользуемся.

Смысл теоремы умножения вероятностей состоит в том, что она часто облегчает введение безусловных вероятностей, если они сразу не очевидны. Приведем соответствующий *пример*.

Рассмотрим опять задачу о раздаче экзаменационных билетов из § 1. Предположим, что номера экзаменационных билетов, доставшихся студентам, для нас ненаблюдаемы, а наблюдаемо лишь выражение лица — счастливое или несчастное. Допустим, что мы верим в принцип, согласно которому ненаблюдаемые вещи не следует вводить в модель. Тогда для ситуации с первыми двумя студентами у нас будут 4 элементарных события: 1с 2с (первый счастливый, второй счастливый), 1с 2н (первый счастливый, второй несчастный) и аналогично 1н 2с, 1н 2н. Как найти вероятность $\mathbf{P}(1с\ 2с)$? Запишем теорему умножения вероятностей

$$\mathbf{P}(1с\ 2с) = \mathbf{P}(1с)\ \mathbf{P}(2с/1с)$$

и с удовольствием убедимся, что входящие в правую часть вероятности нам совершенно ясны:

$$\mathbf{P}(1с) = \frac{n}{N}, \quad \mathbf{P}(2с/1с) = \frac{n-1}{N-1}.$$

Поскольку $2с = 1с2с \cup 1н2с$, аналогично получаем

$$\mathbf{P}(2с) = \mathbf{P}(1с\ 2с) + \mathbf{P}(1н\ 2с) = \frac{n}{N} \cdot \frac{n-1}{N-1} + \frac{N-n}{N} \cdot \frac{n}{N-1} = \frac{n}{N},$$

т. е. шансы первого и второго одинаковы.

Конечно, математически аккуратное решение этой задачи должно включать определение безусловных вероятностей всех четырех элементарных событий и проверку единственной аксиомы (сумма вероятностей элементарных событий равна 1). Не мешает убедиться и в том, что условные вероятности, найденные из безусловных с помощью математического определения, совпадают с теми и без того ясными условными вероятностями, которые послужили опорой нашей интуиции. За математическую строгость надо платить, совершая иногда лишние действия (лишние — с точки зрения здравого смысла). Другим важным неудобством предложенного сейчас решения является его крайняя громоздкость для случая более чем двух студентов. Таким образом, введение только мыслимых, но ненаблюдаемых событий, сводящее всю задачу к частному случаю групповой инвариантности (перестановка студентов в очереди эквивалентна перестановке билетов в куче, разложенной на столе; однако все порядки билетов в куче равновероятны), является предпочтительным.

С условными вероятностями связаны две формулы исчисления, с которыми необходимо познакомиться.

Формула полной вероятности. Пусть множество элементарных событий Ω разбито на n непересекающихся подмножеств $H_1, \ldots, H_n$:

$$\Omega = H_1 + H_2 + \ldots + H_n.$$

Пусть $B \subseteq \Omega$. Тогда имеем $B = BH_1 + BH_2 + \ldots + BH_n$;

$$\mathrm{P}(B) = \mathrm{P}(BH_1 + \ . \ . \ . + BH_n) = \sum_{i=1}^{n} \mathrm{P}(BH_i) =$$

$$= \sum_{i=1}^{n} \mathrm{P}(H_i)\mathrm{P}(B/H_i).$$

Эта формула называется формулой полной вероятности*. Подмножества $H_1, H_2, \ldots, H_n$ называются *гипотезами*. Полученная формула в виде «заклинания» выглядит так: «вероятность какого-нибудь события равна сумме вероятностей гипотез, умноженных на условную вероятность события при данной гипотезе».

Формула Байеса. Произведем выкладку

$$\mathrm{P}(H_i/B) = \frac{\mathrm{P}(H_iB)}{\mathrm{P}(B)} = \frac{\mathrm{P}(H_i)\mathrm{P}(B/H_i)}{\sum\limits_{i=1}^{n} \mathrm{P}(H_i)\mathrm{P}(B/H_i)}.$$

Полученная формула называется формулой Байеса. В этой формуле вероятности гипотез $\mathrm{P}(H_i)$ называются ***априорными***, а условные вероятности $\mathrm{P}(H_i/B)$ — ***апостериорными***. Эти названия связаны со следующей схемой научного исследования.

О природе некоторого явления имеется n гипотез $H_1, H_2, \ldots, H_n$, в которые мы верим с вероятностями $\mathrm{P}(H_1)$, $\mathrm{P}(H_2)$,

* Формула полной вероятности, очевидно, справедлива в случае не только конечного, но и счетного разбиения $\Omega = H_1 + H_2 + \ldots$.