ИСТОРИЯ НАУКИ И ОБУЧЕНИЕ НАУКЕ

(НА ПРИМЕРЕ ПОНЯТИЙ «СЛУЧАЙНОСТЬ» И «ВЕРОЯТНОСТЬ»)

Ю. В. ЧАЙКОВСКИЙ

В предыдущей статье [1] на аналогичную тему уже говорилось, что изучение науки, производимое в свете ее истории, поможет выявить ее концептуальный каркас, избежать догматизма и обманчивого впечатления ее завершенности. К сожалению, многие читатели сочли статью попросту отрицательной рецензией на соответствующие пособия, хотя вместо них можно бы взять и другие. Почти всюду история служит лишь как дидактический материал, освещающий (и освящающий) ныне господствующие взгляды; такая история мало кому полезна. Нужна другая, осмысливающая история, т. е. наука, показывающая прошлое в целом (а не только путь к «сегодня»), раскрывающая место нынешнего «переднего края» в системе знаний.

Указанное несоответствие между наличной и желаемой историей — достаточно общее явление, искажающее восприятие многих наук. Не зная истории своей науки, самые разные ученые оказываются в плену нынешней моды и способны десятилетиями следовать очень старым, архаичным познавательным схемам, воспринимая чисто технические результаты как принципиальные достижения и в то же время отказываясь обсуждать концептуальные вопросы как «ненаучные». В биологии, например, «ненаучным» до самого последнего времени считался вопрос о собственных, не зависящих от отбора, определенных законах изменчивости, а в математике — о природе случайности. Но если наука перестает анализировать свои основы, она понемногу превращается, по выражению методолога Т. Куна, в решение головоломок. Сейчас, когда респектабельная наука признаёт, наконец, определенную изменчивость [2], пора напомнить, что тем самым в биологии меняется и взгляд на случайность: мы не можем теперь строить математическую модель наследственности на утверждениях вроде «допустим, что мутация происходит случайно с вероятностью p» — такие допущения надо обосновывать либо отвергать. Где прочесть об этом? В руководствах и справочниках подобные вопросы даже не ставятся.

Обоснование феномена случайности выпало из круга интересов математики в начале XIX в., когда математика случайного получила название «теория вероятностей» (благодаря книге П.-С. Лапласа [3]; до этого в заглавия книг выносились термины *предположение, догадка, шанс, случай, среднее,* но не *вероятность)*. Именно тогда понятие вероятности стало исходным и вопросы вроде «возможна ли случайность, не обладающая вероятностью?» фактически перестали быть математическими.

Природой случайности продолжают (и с успехом) заниматься философы, но отсутствие у них должного контакта с математиками наносит огромный ущерб науке. Выход видится мне в том, что само преподавание математики случайного должно строиться на вразумительном анализе феномена случайности. А это, как я постараюсь показать, короче всего получается, если вспомнить историю. Чтобы подчеркнуть параллель с работой [1], материал собран в параграфы, носящие те же, что в [1], названия.

Многословная пустота

Энциклопедии [4, 5] не содержат термина «случайность», упоминая лишь «случайное событие» теории вероятностей, причем энциклопедия [5] наиболее радикальна: «Случайное событие — любая комбинация исходов некоторого опыта, имеющая определенную вероятность наступления» [5, т. 5, с. 19]. Вот так. А если определенной вероятности нет, то событие нельзя называть случайным? Как же быть с событиями, об условиях наступления которых просто ничего не известно? Ведь именно их мы чаще всего (в том числе в точных науках) аттестуем как случайные.

Обе энциклопедии определяют вероятность как «численную характеристику возможности наступления события в тех или иных определенных, могущих повторяться неограниченное число раз условиях». Само по себе это верно, но разве на практике мы отказываемся от математического описания события, условия которого реализуются лишь несколько раз? И главное: разве неограниченное повторение условий неизбежно влечет наличие вероятности? Нет, разумеется, тут дело в конкретном опыте. «Как показывает практика, при больших N частоты $N(A)\,/\,N$ в различных сериях испытаний оказываются приблизительно одинаковыми [1], группируясь около некоторого числа $p\,(A)$, которое и называют вероятностью события A» [4, т. 1, с. 184]. Однако бывает и другая «практика»: возьмите наугад несколько русских книг и подсчитайте частоту употребления каждого слова в каждой — частоты редких слов будут совсем различны. Можно взять миллион книг, но частоты к пределам не сойдутся — наоборот, не так уж редко будут появляться все новые и новые слова, хотя самих слов в русском языке меньше 200 тысяч, а в опыте вашем будет более 200 миллиардов словоупотреблений. Частоты в подобных опытах неустойчивы, но руководства этот факт игнорируют.

Каковы же те пределы, в которых вероятность существует теоретически? А в каких — практически? Или: чем теория вероятностей занимается, а от чего отказывается? Вот ответ: «Вероятностей теория — математическая наука, позволяющая по вероятностям одних случайных событий находить вероятность других» [5, т. 1, с. 655]. Это значит, что теория вероятностей — лишь малая часть науки о случайном. Если не помнить этого, то «важнейший вопрос о применимости теоретико-вероятностных понятий... полностью выпадает из поля зрения» [6, с. 16]. Применимость — вопрос, относящийся к реальной природе, поэтому Дж. Литлвуд писал: «Смысл вероятности должен иметь отношение к реальному миру, и должно существовать... первичное предположение, «аксиома вероятности», верная в естественно-научном (а не в абстрактном) смысле [7, с. 59, 61]. Таково мнение математиков.

Обратимся теперь к философам: «Случайность — отражение в основном внешних, несущественных, неустойчивых, единичных связей действительности; выражение начального пункта познания объекта; результат перекрещивания независимых причинных процессов...» [8, с. 421]. Как и следовало ожидать, подход здесь куда шире и глубже, но как только речь заходит о связи случайности с вероятностью, начинается путаница, например: «Классическая интерпретация вероятности уступила место статистической» [8, с. 79]. Классическая интерпретация (если у кости M граней, то вероятность выпадения данной грани равна $1/M$) не может «уступить место» статистической, поскольку они взаимодополнительны: первая — абстракция, вторая — опыт. В действительности классическая интерпретация уступила место аксиоматической, но сами математики настолько обходят вопрос о связи статистического подхода с прочими (для проверки можно взять любой учебник, даже лучший — курс Вильяма Феллера [9]), что винить философов в путанице нельзя.

---

[1] Здесь N — число испытаний, $N(A)$ — число наступлений события A. Такое определение вероятности называют *статистическим*.

Само изложение теории у Феллера превосходно, но тем досаднее многословная пустота вводных глав. По-видимому, они любопытны для психолога — как свидетельство смущения учителя, неспособного объяснить ученику исходные понятия, но для вхождения в предмет не дают ничего. Суть дела исчерпывается поразительной фразой: «Вероятности... заданы и не нуждаются ни в каких предположениях об их действительном численном значении или о способе их измерения на практике» (с. 37) и столь же поразительной уверенностью, что читатель сам увидит в вероятности «наглядный смысл» (с. 38). Разве дело в измерении или в наглядности? Сперва надо убедиться, что есть объект для измерения, а об этом нет и речи.

Объект для измерения (вероятность) есть там, где есть устойчивые частоты наступления однотипных событий, но сама однотипность — простейшая ситуация, грубая идеализация (подробнее см. [10]). В курсе Б. В. Гнеденко читаем: «... утверждение, что эта вероятность существует, является содержательным утверждением, нуждающимся в каждом отдельном случае в обосновании» [11, с. 17], но и тут обоснований нет. Почему?

Дело в том, что вопрос о связи частоты с вероятностью совсем непрост, он уходит в глубь проблемы о природе случайного. Как писал А. Н. Колмогоров, «само понятие математической вероятности было бы бесплодно, если бы не находило своего осуществления в виде частоты появления какого-либо результата при многократном повторении однородных условий. Поэтому... настоящая ее (теории вероятностей.— *Ю. Ч.*) история начинается с закона больших чисел Я. Бернулли» [12, с. 4]. Если так, то закон больших чисел (ЗБЧ) и есть искомая связь природы с теорией, т. е. надо обратиться к Бернулли.

В энциклопедии читаем: Якоб Бернулли «доказал теорему, утверждающую, что в последовательности независимых испытаний, в каждом из которых вероятность наступления некоторого события А имеет одно и то же значение p, $0<p<1$, верно соотношение

$$P\{|\mu_n / n - p| > \varepsilon\} \to 0 \qquad (1)$$

при любом $\varepsilon > 0$ и $n \to \infty$; здесь μ_n — число появления события А в первых n испытаниях». Вопреки обычному лаконизму энциклопедий, это сказано — и с массой подробностей — дважды [5, т. 1, с. 425, 523], ЗБЧ описан и как закон природы, и как основной результат теории вероятностей, но не сказано главного — как наука вычислять «по одним вероятностям другие» смогла что-то сказать о связи вероятности с опытом, с частотой. Опять многословная пустота. Примерно то же — в других руководствах.

Вместо доказательства существования предела p его наличие постулируется. В такой редакции у ЗБЧ остается весьма скромный смысл — лишь оценка скорости сходимости. Определение вероятности p через вероятность P не раз аттестовали как порочный круг. Потребитель же ждет от теории, прежде всего указания на границы того класса явлений природы, в котором частоты устойчивы, а затем уже численных оценок сходимости. «Действительно важной задачей является не формальное уточнение этого определения (вероятности через ЗБЧ.— *Ю. Ч.*), а возможно более широкое выяснение условий, при которых такого типа вероятностная случайность должна проявляться»,— писал Колмогоров [13, с. 275]. Именно такого выяснения мы нигде не видим, да и сам Колмогоров признавал устойчивость частот не математическим, а природным феноменом, «тенденцией» [13, с. 270, 275]. Есть точка зрения, что логическое обоснование границ применимости вероятностных методов вообще недостижимо [14, с. 57]. При таком подходе ЗБЧ оказывается никак не основой теории, не связью случайности с вероятностью, а всего лишь оценкой скорости сходимости частоты к пределу, если предел существует. Но как же тогда на деле природа случайного связана с теорией случайного, с теорией вероятностей? Ответа нет, и недаром почти 300 лет шел спор, считать ли теорию вероятностей математической наукой.

Принято думать, что спор решен появлением первой аксиоматики Колмогорова (1933), но, по-моему, здесь произошло то же, что с дарвинизмом: снятие одной частной трудности воспринято как окончательное решение проблем обоснования вообще. В эволюционизме есть две связанные проблемы: как изменяется признак и как создается организация. Когда в начале века было решено (как сейчас выяснилось, ошибочно), что признак меняется путем мутации, воцарилось убеждение, что эволюционизм обоснован как таковой. Аналогично, в теории вероятностей есть две связанные проблемы: что считать элементарным событием и как частота порождает вероятность. На эту связь настойчиво указывал Рихард Мизес (1931), и он же решил первую проблему, предложив понятие пространства элементарных событий, которое вскоре было включено в первую аксиоматику Колмогорова, о чем сказано у Феллера [9] (подробнее это сделано в первом издании [15]). Что же касается второй проблемы (проблемы смысла ЗБЧ), то вновь и вновь предлагаются различные подходы [16, 17], и споры не затухают. По Феллеру, дискуссия касалась только второстепенных вопросов, интересных лишь немногим специалистам» [15, с. 14], причем из следующего издания Феллер изъял даже и эту ремарку, но вопрос не стал от этого яснее. По-моему, проблема касается всех, кроме, быть может, тех, кто вычисляет «по одним вероятностям другие». Однако, восприняв ее при обучении как решенную, подавляющее большинство к ней уже никогда не может вернуться.

Разумеется, студенты-вероятники в ходе обучения понимают, что о вероятности следует говорить только там, где она имеет место, а об остальном им (вероятникам) говорить не следует. Отсюда и определение (нелепое для прикладника [6, с. 15]) случайного события как чего-то, имеющего вероятность. Наблюдаемые в реальном мире частоты далеко не всегда устойчивы, и, прежде чем учить вычислению «по одним вероятностям других», хорошо бы рассказать, что такое случайность и как она связана с вероятностью. Простейший путь здесь, по-моему,— кратко освещать историю становления этих понятий. Она известна нескольким историкам (пишущим в основном по-английски), но научным сообществом забыта, и это приводит к огромным издержкам, вплоть до утраты смысла в целом ряде работ научного и прикладного характера.

Вплоть до утраты смысла

Формула (1) приводится всюду как формула Бернулли, но в действительности она на 150 лет моложе самой теоремы Бернулли и принадлежит П.-С. Пуассону (1837). Бернулли еще не знал понятий, нужных для ее записи, и доказал другое утверждение, лишенное того порочного круга, который ему приписали позже, когда первоначальный смысл ЗБЧ был утрачен.

Мало найдется в математике утверждений, о которых высказано так много противоречащих друг другу мнений, как об этой, доказанной около 1690 г. и опубликованной в 1713 г. теореме Бернулли. Один из лучших знатоков проблемы, Ян Хэкинг, полагает, что теорема обеспечила автору «место в пантеоне вероятников» [18, с. 154], иногда ее даже называют «золотой теоремой» [19, с. 136, 138]; однако не менее компетентный автор счел ее лишь занимательной задачей комбинаторики с довольно бедным математическим содержанием [20, с. 32]. Одни говорили, что она точна и строга [12, с. 19], другие — что в основе здесь порочный круг, лишающий теорему всякой приложимости [21, с. 85], третьи — что она «играет первостепенную роль в различных практических применениях» [22, с. 98], четвертые — что былое увлечение теоремами такого типа «обусловливалось не вполне логичной их интерпретацией» [6, с. 27]. Если для А. А. Маркова само доказательство «имеет свои достоинства и до сих пор не потеряло интереса» [12, с. 19], то его современник К. Пирсон счел доказательство теоремы никуда не годным [18, с. 154; 22, с. 100], а для Э. Бореля с соавторами она вообще «не является доказательством» [23, с. 59]. Наконец, одни видели здесь только классическое определение вероятности, другие — только статистическое [22, с. 96]. Обще, пожалуй, только одно: не анализируется отличие формулы (1) от формулировки самого Бернулли.

Вот она: «Пусть число благоприятных случаев относится к числу неблагоприятных точно или приближенно как r к s или к числу всех случаев — как r к $r+s$ или r к t, каковое отношение заключается в пределах $(r+1)/t$ и $(r-1)/t$. Требуется доказать, что можно взять столько опытов, чтобы в какое угодно данное число раз (c раз) было вероятнее, что число благоприятных наблюдений попадет в эти пределы, а не вне их, т. е. что отношение числа благоприятных наблюдений к числу всех будет не более, чем $(r+1)/t$, и не менее, чем $(r-1)/t$» [12, с. 56].

Ясно видно, что никаких «независимых испытаний» здесь нет. Более того, здесь в отличие от позднейших формулировок только однажды упоминается вероятность — в слове «вероятнее»; определения вероятности p через вероятность P, какое мы видим в (1), здесь тоже нет. Наконец, если бы не слово «вероятнее», то сама теорема имела бы чисто арифметическую формулировку, и, оказывается, от этого слова вполне можно здесь избавиться, поскольку оно не нужно при доказательстве. Если проследить за доказательством, то выяснится, что автор сравнивал вовсе не случайные события, а некоторые наборы из разного числа одинаковых элементов и называл «более вероятным» тот набор, в котором элементов больше. Следовательно, теорема допускает чисто арифметическую (точнее, комбинаторную) формулировку, которую стоит здесь дать.

Даны N одинаковых симметричных t-гранных игральных костей, грани которых перенумерованы — от 1 до t на каждой кости. (Тот факт, что существует всего 5 типов правильных t-гранников ($t=4, 6, 8, 12, 20$), не играет роли, так как симметричную кость можно представить, например, в виде двух t-гранных пирамид, склеенных основаниями; случай $t=2$ соответствует симметричной монете.) Разложим кости в ряд на столе и отметим, на какой грани лежит каждая. Различных способов раскладки костей будет t^N. Закрасим на каждой кости r граней — число способов раскладки не изменится, но теперь про каждую кость можно еще вдобавок отметить, лежит ли она на закрашенной грани или нет. (В вероятностных терминах В. П. Криндача [24, с. 263], номер грани — микросостояние, а совокупность закрашенных граней — макросостояние. Микросостояние в старой литературе называли шансом или статочностью [25, с. 11], а макросостояние — точнее, долю данного макросостояния в общей совокупности состояний — чаще именуют вероятностью в классическом смысле). Теорема Бернулли утверждает: если N достаточно велико, то в подавляющем большинстве раскладок окажется, что доля костей, лежащих на закрашенной грани, близка к $^r/_t$.

Доказательство, хоть и длинное, в сущности просто. Сперва t^N записывается в виде $(r+s)^N$ и раскладывается на слагаемые по формуле бинома Ньютона. В этой сумме отыскивается затем самое большое слагаемое — им оказывается *центральный* член, соответствующий раскладке, для которой доля закрашенных нижних граней равна r/t в точности. Бернулли нашел, что при достаточно большом N почти вся сумма приходится на это и несколько соседних слагаемых, причем эти слагаемые означают количество тех раскладок, в которых почти rN/t костей лежат на закрашенных гранях. Итак, доля закрашенных оснований близка к r/t почти во всех мыслимых раскладках.

Вот все, что доказано у Бернулли, остальное — интерпретации. Первую дал сам Бернулли, и ее часто цитируют: «Если бы наблюдения над всеми событиями продолжать всю вечность (при чем вероятность, наконец, перешла бы в полную достоверность), то было бы замечено, что все в мире управляется точными отношениями и постоянным законом изменений, так что даже в вещах, в высшей степени случайных, мы принуждены были бы признать как бы некоторую необходимость и, скажу я, рок» [12, с. 59]. Другими словами, автор видел в своей арифметической теореме указание на некий всеобъемлющий закон природы, а в самой природе — господство *равновозможностей*: если все раскладки равноправны, то любая из них должна встречаться одинаково часто. Бросив N костей на стол как попало, мы обязательно получим одну из раскладок, уже встречавшуюся при методическом раскладывании; если же все раскладки равновозможны, то мы будем обнаруживать долю, близкую к r/t, почти всегда. Давая этой игровой схеме столь общую интерпретацию («наблюдения над всеми событиями»), Бернулли прямо писал: если события сами по себе не обладают равновозможностью, то «необходимо уравнять их и вместо каждого легче встречающегося случая считать столько других, насколько он легче имеет место» (с. 34). Эта процедура — приведение любой ситуации к набору равновозможных случаев и взятие каждого случая ровно один раз — и есть искомая связь случайности с вероятностью, «аксиома вероятности» по Литлвуду.

Как видно, никакой случайности в схеме рассуждений Бернулли просто нет: вместо интуитивно приемлемого убеждения, что равновозможные события происходят одинаково часто в каком-то среднем смысле, они здесь попросту взяты все по одному разу. В этом — равновозможность отдельных исходов. Аналогично, взяв ровно по одному разу каждую мыслимую серию испытаний (раскладку), Бернулли ввел (или, если угодно, обошел) независимость испытаний [2] Это может показаться примитивом, но именно данная идея легла через 200 лет в основу понимания вероятности как меры (см. § 3), и именно она позволила теории вероятностей описывать не только азартные игры, но и устойчивые случайности вообще. В самом деле, обычно понятие классической вероятности иллюстрируется в учебниках на азартных играх потому, что здесь равновозможность прямо задана внешней, геометрической симметрией объекта — монеты, игральной кости, колоды карт и т. п. Однако у Бернулли равновозможность имеет куда более широкий смысл: если монета симметрична, то ее следует рассматривать как «двугранную кость», если же она почти симметрична, то — как симметричную t-гранную кость, у которой $r \approx t/2$,— вот и вся разница. Равновозможность здесь — не тот редкий частный случай, каким оперируют учебники, а единица измерения, как бы общий делитель. В качестве постулируемого, далее не анализируемого свойства природы взята никак не внешняя симметрия монеты, а та внутренняя симметрия пространства-времени, которая позволяет моделировать все бросания данной монеты посредством одной t-гранной кости. О связи вероятности и симметрии см. [24, с. 256—267].

Симметричные формулировки стали возможны после Пьера Кюри, который в 1884 г. сформулировал идею симметрии как основы природных явлений [24, с. 13, 181]. Еще в начале XX в., когда Эмми Нётер уже положила принцип симметрии в основу механики, для теории вероятностей он оставался чуждым. Многие (более всех Мизес [21]) считали основой схемы Бернулли тавтологию (вероятность определяется через равновозможности, а они рассматриваются как равные вероятности). В действительности здесь тавтологии не больше, чем в таком, например, определении: события называются независимыми, если вероятность их совместного наступления равна произведению их вероятностей [9, с. 143; 20, с. 25]. Это никого не шокирует: хотя «мы на самом деле не знаем, что такое независимость, но чем бы она ни была, если она имеет смысл, то она должна обладать» этим свойством [26, с. 72]. Точно так же можно сказать: мы не знаем, что такое равновозможность, но если она имеет смысл, то ею должны обладать грани симметричной кости.

И вот парадокс: хотя уже полвека принято упрекать Мизеса в непонимании природы вероятности, хотя аксиоматика (а значит, и аппарат) строится по Колмогорову, однако концепция применений следует всегда Мизесу [14, с. 15], т. е. игнорирует феномен симметрии природы и апеллирует исключительно к эмпирической частоте. Издержки такого бездумия огромны: ведь статистически частоту и другие средние величины можно измерить всегда, при любом повторяющемся явлении, даже когда они неустойчивы от серии к серии, т. е. когда вероятностей

[2] Независимость испытаний — лишь позднейшая интерпретация теоремы Бернулли: кроме случайных независимых теореме удовлетворяют и строго детерминированные процессы (перебор раскладок).

нет. Так в биологии произошло при построении «синтетической теории эволюции», в основе которой постулирован сдвиг средних значений наблюдаемых признаков в сторону большей приспособленности, хотя каждый из этих признаков обычно описывается гиперболическим распределением, не имеющим среднего значения [27, 28]. И вряд ли можно винить в этом биологов, поскольку гиперболические распределения до сих пор едва упоминаются в учебниках статистики и в [5]. А ведь они наглядный и очень обычный пример массовых процессов, в которых ЗБЧ на выборках приемлемого размера не выполняется.

Зная подлинную теорему Бернулли, понять причину нарушения ЗБЧ просто: он нарушается там, где случайное явление нельзя представить с помощью единой t-гранной кости ни при каком ограниченном t, т. е. когда множество мыслимых исходов быстро растет в ходе испытаний. Простейший пример — размножение бактерий: если каждая бактерия за час либо гибнет, либо (с той же вероятностью) делится на две, то через T часов среднее количество бактерий не изменится, но количество наблюдаемых особей может оказаться от 0 до 2_T, т. е. для моделирования нужна 2^T-гранная кость, причем $T \to \infty$. В этой ситуации среднее не несет реального смысла: во-первых, для оценки среднего надо провести более 10 T^2 независимых опытов, что нереально, а во-вторых, реальное количество не приближается к среднему — в каждом опыте оно обычно либо быстро достигает нуля, либо быстро растет много выше среднего [29]. Это азы теории ветвящихся процессов, разрабатываемой более 100 лет и хорошо изложенной, например, у Феллера. Однако тот факт, что идеология ветвящегося процесса противоположна идеологии ЗБЧ, узнать до сих пор негде и нерегулярность своих опытов микробиологи до сих пор целиком объясняют «грязью».

Почему так? Для ответа следует выявить основную традицию развития науки о случайном. Здесь-то работа историка только и начинается.

На кого работает историк науки

Устойчивость частот фактически сознавали уже средневековые банкиры, но они тщательно скрывали секреты своей науки, и мы судим о ней только по ее результатам — успешной деятельности некоторых банков в неимоверно сложных условиях [30]. Недаром великий ученый XVI в. Джироламо Кардано сопоставлял финансовую науку с черной магией. Сам он около 1550 г. написал труд «Об азартных играх», увидевший свет только в 1663 г.,— первую работу по теории вероятностей, как мы ее сейчас понимаем. До него об азартных играх писали исключительно в комбинаторном плане, т. е. в смысле вычисления «по одним вероятностям других». Эта линия рассуждений привела в середине XVII в. к знаменитым выводам Б. Паскаля и П. Ферма, с которых обычно, забывая о Кардано, начинают историю теории вероятностей. Она хорошо известна [18, 19, 22], и мы отметим лишь одно — Паскаль прямо противопоставил задачи новой науки задачам статистического эксперимента: «Колебания счастья и удачи подчиняются рассуждениям, опирающимся на справедливость... Это в тем большей мере должно определяться усилиями разума, чем в меньшей мере может быть найдено из опыта. Ведь неопределенный исход явления теснее связан со случайностью, чем с законами природы» [25, с. 95]. Естественно, что «справедливость» понималась как равновозможность, а сама равновозможность имела более чем широкий смысл.

Во времена Кардано игроки использовали заведомо несимметричные кости в качестве симметричных, да и много позже (когда уже существовала наука о случайном) равновозможность подчас входила в расчеты без всякого обсуждения. Так, великий натуралист Ж.-Л. Бюффон в 1749 г. завершил свой труд о человеке обширной таблицей, содержавшей «вероятности» человеку данного возраста умереть в течение суток, трех месяцев, года и т. д. [31]. Мало того, что выводы делались на совсем небольших выборках (так, для последнего возраста — 98 лет — Бюффон располагал данными всего о 8 людях), но и сама вероятность умереть, не зависящая от внешней обстановки, пола, здоровья и рода занятий, не могла быть особо полезной. Зато она хорошо иллюстрирует идейную подоплеку равновозможности: все, что трудно или нельзя определить, проще всего считать несущественным. Впоследствии эта идея была даже эксплицирована в форме принципа индифферентности, оказавшего теории вероятностей плохую услугу [16, с. 47].

В противовес идее равновозможности развивался чисто статистический подход к вероятности, перешедший от банкиров к статистикам и проникший прежде всего в демографию. Замечательные мысли на этот счет мы видим у Джона Граунта (1662), который первым сумел [32] чисто интуитивно отличать реальные тенденции движения населения от случайных вариаций. Его работы широко обсуждались, и некоторые ученые (Х. Гюйгенс, Г.-В. Лейбниц и др.) пытались увязать их с академическими упражнениями Паскаля — Ферма, но найти связь вероятности-равновозможности с вероятностью-частотой долго не удавалось.

А ведь об этом ясно писал еще Кардано. Отметив, что при игре в две кости есть 6 возможностей выпадения пары одинаковых очков, что составляет 1/6 часть от общего числа (36) возможных здесь пар, он заметил: «При большом числе игр оказывается, что действительность весьма приближается к этому предположению», т. е. к тому, что каждый шестой исход — пара равных очков. В другом месте он уточнил: не следует думать, что, бросив кости много раз, мы получим точно требуемое число очков, «но при бесконечном числе бросков это должно случиться почти обязательно, ибо в большом числе повторений проявляется течение времени, демонстрирующее все формы» [33, с. 265, 267].

Трудно толковать мыслителя прошлого, но все же в этой фразе явственно звучит нечто дьявольское. Только Якобу Бернулли дано было лет через 150 после ее написания уточнить смысл. Ссылки на Кардано у него нет, но развита именно карданова мысль: Бернулли в явном виде «продемонстрировал все формы» и учел их в своем выводе *ровно по одному разу* — до него это делали только с отдельными исходами, он же брал доли исходов, а затем и серии исходов (раскладки). Он снял, в понятиях своей эпохи, карданово «почти», показав (если применить наши нынешние термины), что P в формуле (1) стремится к нулю (а не к иной «асимптоте» [12, с. 42]). Однако прошло еще 200 лет, и Эмиль Борель как бы вернулся к этому «почти», показав (1909), что оно имеет глубокий смысл в контексте теории меры Лебега: «почти обязательно» значит теперь, что бесконечные серии, в которых частота не равна вероятности, возможны, но имеют меру нуль [20]. При конструировании вероятности как меры каждую точку измеримого множества тоже, как у Бернулли, берут, ровно один раз, т. е. случайность здесь тоже не столько описывают, сколько обходят — постольку, поскольку это не ведет к противоречию с опытом. Такова теория вероятностей по Колмогорову (точнее, первая аксиоматика Колмогорова, 1933). Эту традицию назовем *исчерпанием равновозможностей.* Наконец, карданово «течение времени» ассоциируется ныне с однородностью событийного пространства-времени, т. е. равновозможностью тех формальных исходов (микросостояний), на которые Бернулли так удачно разложил исходы реальные (макросостояния). Однородность событийного пространства-времени — феномен, который еще предстоит изучить. Возможно, что именно он ответствен за устойчивость частот [28, с. 104]. В таком случае это одна из тех фундаментальных симметрий, которыми определяются законы сохранения; в физике их выявила теорема Эмми Нётер. Вот сколько вех приходится вспоминать при чтении одной фразы Кардано.

Кроме комбинаторного (априорного) и статистического (опытного) в XVII в. развивалось понимание вероятности как степени уверенности (моральная вероятность), о чем можно прочесть у Хэкинга [18] и в комментарии О. Б. Шейнина к работе [12]. Все три понимания были синтезированы Якобом Бернулли в его трактовке термина «вероятность», поэтому-то историки и находят у него то, что им созвучно. Обсудить все это здесь невозможно, но надо заметить, что, насколько мне известно, никогда впоследствии эти три понимания воедино не сводились (литературу см. [16—18]). Поэтому содержательное обучение лучше начинать с Кардано и Бернулли, а не с комбинаторики Паскаля и не с «теоремы Бернулли» в ее нынешнем толковании. После этого становится само по себе ясно, что теория вероятностей в понимании Лапласа или Колмогорова является развитием первого (априорного) понимания, а вовсе не второго. Второе (опытное) лишь по традиции фигурирует во введениях к учебникам, чтобы потом исчезнуть из текста.

Поразительно, но даже в математической статистике обычно разрабатывается не статистическое, а априорное понимание вероятности. Так, Г. Крамер, автор «Математических методов статистики» (русский перевод: 1948), считал вероятность всего лишь «абстрактным двойником» наблюдаемой частоты, и такое понимание стало общим, причем без дальнейшего анализа [16, с. 68, 76]. Собственно статистическое понимание вероятности прозябает на периферии научной мысли и связывается с именем Мизеса [21], хотя основная его идея была известна и ранее [18, с. 54]. Статистическое понимание состоит в том, что постулируется не симметрия (равновозможность), а беспорядочность. При этом возникают трудности: в точности показать, в каких случаях частота беспорядочных событий сходится к пределу (вероятности), а в каких — нет, до сих пор не удавалось, хотя последователи Мизеса достигли определенного успеха — в некотором усеченном смысле понимать вероятность как предел частоты можно [16, гл. 4]. Однако было бы ошибкой полагать, что найдена какая-то альтернатива равновозможности. По-видимому, без равновозможности нет и вероятности, но это не только не доказано, но и сама задача о доказательстве, насколько мне известно, точно не поставлена. Колмогоров ограничивался указанием, что среди случайных явлений есть «широкий класс явлений», обладающих вероятностями, и называл их вероятностно-случайными, или *стохастическими* [13, с. 254]; другие авторы обычно не делают даже этого, и стохастичность понемногу стала восприниматься как синоним случайности.

На мой взгляд, главный просчет Мизеса состоял в отказе от симметрии, от того подхода, который к тому времени реализовали в физике Кюри и Нётер, т. е. от понимания вероятности как инварианта, взаимодополнительного к пониманию ее как предела. Зато именно Мизес привлек общее внимание к сути феномена беспорядка, который игнорируется идеологией исчерпания равновозможностей. Эта идеология вовсе не игнорирует опыт и, даже наоборот, она когда-то черпала аргументацию в опыте. Историк Вильям Адамс напоминает, что в XVII в. распространился прием уточнения измерений путем использования среднего арифметического [34, с. 33]. (До этого все, даже астрономы, предпочитали выбирать «более надежное» измерение, т. е. поступали примерно так же, как до сих пор поступают микробиологи.) Большой надежностью среднего пользовался, например, Граунт [32, с. 31, 72 и др.]. В 1755 г. Томас Симпсон поставил вопрос о математическом обосновании того факта, что среднее точнее отдельного наблюдения [34, с. 34]. Этот опытный факт толковался с позиций симметрии (отклонение от истинного значения вправо происходит при измерении столь же часто, сколь и такое же отклонение влево), что позволило использовать идеологию ЗБЧ, и с этого соображения началась история *нормального закона ошибок*.

Нормальное распределение случайной величины исследовалось начиная с 1733 г., с работ Абраама де Муавра. Муавр выразил факториалы в биноме Ньютона формулой $n! \approx \mathrm{C} n^{n} \sqrt{n} \cdot e^{-n}$, а Джеймс Стирлинг вскоре нашел, что $\mathrm{C} = \sqrt{2\pi}$, и вся формула почему-то носит его имя [34, с. 21]. Затем Муавр аппроксимировал сумму k членов бинома Ньютона длины n в теореме Бернулли с помощью интеграла $\sqrt{2/\pi n} \int_0^k \exp(-2t^2/n)\,dt$, откуда и вывел (положив $2k = \sqrt{n}$) свою знаменитую оценку: $\sqrt{n}$ есть мера разброса случайной величины при n измерениях [34, с. 24]. Казалось бы, симметрийная природа нормального закона вполне ясна — она идет от ЗБЧ, однако закон оставался загадочным, и еще в начале XX в. Анри Пуанкаре говорил полушутя, что в нем должно быть что-то таинственное, поскольку математики называют его законом природы, а физики — математической теоремой [20, с. 80]. Легко понять, что вся «тайна» — в утере первичного смысла ЗБЧ. Бернулли как бы постулировал симметрию случайного, и она действительно в огромном классе событий имеет место.

За 80 лет положение не прояснилось. До сих пор математики пишут, что приложимость нормального закона к ошибкам измерений «обусловлена тем, что он превосходно согласуется с фактами» [23, с. 142]. Популяризаторы, естественно, идут еще дальше: столкнувшись с необходимостью объяснить школьникам, почему данная физическая величина распределена нормально, физик простодушно пишет, что ее нормальность следует из расчетов на ЭВМ [35, с. 31]. Поразительно — природное явление следует из расчетов!

Приходится напомнить: «На самом деле нельзя утверждать, что это [нормальное] распределение возникает непосредственно в физических ситуациях», оно лишь «выступает в качестве предела» [36, с. 200], и причины этого, во-первых, в том, что оно симметрично — обладает самосопряженной плотностью и замкнуто относительно сложения, а во-вторых, в том, что оно проявляет два экстремальных свойства [36, с. 199, 270]. (Первое из них установил в 1809 г. Карл Гаусс.) Поэтому, добавим, нормальное распределение естественно выражает в пределе *тройную симметрию* случайности: разновозможность (событий, серий), аддитивность (когда результирующая случайность — сумма) и одномасштабность (равномерную ограниченность) слагаемых [28, с. 94]. Эта симметрия — свойство природы, которому в теории вероятностей найден «абстрактный двойник» в виде исчерпания равновозможностей.

Другой популяризатор старается быть научнее: «В теории вероятностей доказывается, что сумма достаточно большого числа независимых случайных величин, подчиняющихся каким угодно законам распределения, приближенно описывается нормальным распределением», которое он называет «распределением Лапласа — Гаусса» [37, с. 34, 41]. Тут же вышел прямой обман учеников: известно много типов сумм, не сходящихся к нормальному распределению. Ярким символом неведения служит само выражение «Лапласа — Гаусса»: если бы имена классиков упоминались только теми, кто понял их логику, то Лаплас, развивавший здесь идею Гаусса [34], не оказался бы впереди, да и само распределение, вернее всего, носило бы имена Даниэля Бернулли и Христиана Крампа — ученых, не только понявших до Гаусса роль нормального распределения, но и определивших его численные значения [25, 34, 38]. Главное же: тот, кто понял нормальное распределение, не напишет, что к нему сходятся «какие угодно» суммы. Нарушение тройной симметрии (а с тем и условий центральной предельной теоремы) открывает широкое поле «негауссовой» статистики, чаще всего выступающей в форме гиперболических распределений (где частоты могут быть неустойчивы). Они довольно хорошо обследованы прикладниками у нас (Б. И. Кудрин, В. И. Смирнов, С. Д. Хайтун, Л. Л. Численко, А. И. Яблонский и др.) и за рубежом, но обычно игнорируются при обучении.

Недавно, в 1985 г., я спросил одного профессора-вероятника, автора солидной монографии и прекрасной популярной брошюры: чем он объясняет самому себе нормальность распределения природных величин? Он ответил: интуитивно приемлемо считать реально наблюдаемую величину суммой большого числа малых разнородных влияний, связь между которыми слаба. На мой вопрос, почему разнородные влияния именно суммируются и почему порядок малости их взаимосвязи более высок, чем сами влияния, он ответил, что иначе нормальный закон не наблюдался бы столь широко, и тут же с улыбкой признал здесь порочный круг. Курьезность ситуации тут не в самом обращении к опыту, а в обращении дважды: априори (при постулировании наличия вероятностей, суммирования несовместимых и перемножения независимых) и в ходе конкретного рассуждения, в подгонке равновозможностей под заданный ответ.

Исчерпание равновозможностей — остроумный способ обойти описание случайности как таковой, но ведь надо это описание дать. От Бореля идет (и в 1963 г. формализована второй аксиоматикой Колмогорова) трактовка случайной последовательности как знаков действительного числа: почти всякое из них неконструктивно, т. е. не существует способа вычислить следующий знак числа лишь на основании знания предыдущих (литературу см. [10, с. 155]). По аналогии с этой линией рассуждений и естественно, по-моему, строить связь симметрийного и статистического подходов. Если вероятность не постулируется, а ищется как инвариант системы, то естественно начать с систем, в которых инвариантный (симметрийный) подход разработан,— с динамических. Стохастичность динамических систем исследуется уже 30 лет, и в ее рамках «аксиома вероятности» по Литлвуду оказалась не аксиомой, а доказуемым свойством, которое именуют критерием стохастичности [39]. И если 30 лет мало для попадания в учебники, то остается напомнить, что еще на полвека раньше Пуанкаре описал совсем наглядный пример стохастичности -- *произвольную функцию*. Это понятие тоже игнорируется при обучении, и совсем незаслуженно.

Дан горизонтальный круг, расчерченный на белые и черные секторы, а в центре на вертикальной оси крутится стрелка. Если начальная ее скорость v достаточно велика, то какова бы ни была монотонная непрерывная функция α (v), определяющая угол α остановки, остановка в белом секторе определяется не ею (в этом смысле функция α (v) *произвольна*), а соотношением длин секторов [40, с. 63]. Здесь ясно, откуда берется случайность (нерегулярность): хотя близким v_1 и v_2 соответствуют близкие α_1 и α_2, но они могут принадлежать разным секторам, и случайность порождается разрывностью отображения множества начальных состояний во множество конечных. На этой модели легко пояснить связь двух пониманий вероятности — как частоты остановки в белом секторе и как меры (суммарной площади белых секторов). Видна и связь двух аксиоматик Колмогорова.

Итак, история — хороший ключ к недогматическому обучению науке. Но не панацея: зная внешний ход истории, можно путаться в самой науке — вспомним хотя бы тех, кто увидел у Я. Бернулли только статистический подход (см. выше). Такое «знание» может вредить и преподаванию. Так, учебники до сих пор следуют Муавру и Стирлингу, когда дают длиннейший вывод их формулы. Чаще же вывод просто опускают, что лишает формулу наглядности, снижает ее познавательный смысл и затрудняет усвоение. Возможно, что для будущего вероятника долгий первый путь и правилен, поскольку допускает уточнения формулы, но почему всем другим вообще ничего не объясняют? Ведь главное — насколько факториал растет медленнее, чем степенно-показательная функция — показать очень просто. Положив $n!=n^n a_n$ (где $a_n \to 0$), можно с помощью простейшей втузовской теории пределов в 2 строчки найти: $a_n = b_n\ e^{-n}$, где $b_{n+1}/b_n \to 1$. Столь же легко увидеть, что $b_n = C\sqrt{n}$. Значение константы $C=\sqrt{2\pi}$ можно вывести позже (что учебники часто и делают, тоже очень длинно) и опять в 2 строчки — например, проинтегрировав двумерную нормальную плотность в полярных координатах, где первообразная тривиально выражается элементарной функцией. Словом, стоит пользоваться более простыми, чем 250 лет назад, средствами анализа.

Завершая этот беглый очерк, можно сказать: без историка нарушается целостность знания. Чтобы увидеть связи своей дисциплины с другими, надо уметь видет ее место в исторической панораме. В науке о случайном сейчас безраздельно доминирует теория вероятностей, но круг ее задач ограничен, и именно с формулировки границ следует, по-моему, начинать обучение.

Литература

1. *Чайковский Ю. В.* История науки и обучение науке (на примере дарвинизма) // Вопр. истории естествознания и техники. 1987. № 2. С. 92—101.
2. Мастерство мутирования // В мире науки. 1989. № 1. С. 54—55.

3. *Laplace P. S.* Théorie analytique des probabilités. P., 1812.
4. Энциклопедия кибернетики. Т. 1—2. Киев, 1975.
5. Математическая энциклопедия. Т. 1—5. М., 1977—1985.
6. *Алимов Ю. И.* Альтернатива методу математической статистики. М., 1980.
7. *Литлвуд Дж.* Математическая смесь. М., 1962.
8. Философский энциклопедический словарь. М., 1983.
9. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 1. М., 1984.
10. *Чайковский Ю. В.* Разнообразие и случайность // Методы научного познания и физика. М., 1985. С. 149—168.
11. *Гнеденко Б. В.* Курс теории вероятностей. М., 1961.
12. *Бернулли Я.* О законе больших чисел. М., 1986.
13. *Колмогоров А. Н.* Теория вероятностей // Математика, ее методы и значение. М., 1956. Т. 2. С. 252—284.
14. *Тутубалин В. Н.* Границы применимости (вероятностно-статистические методы и их возможности). М., 1977.
15. *Феллер В.* Введение в теорию вероятностей и ее приложения (дискретные распределения). М., 1952.
16. *Кайберг Г.* Вероятность и индуктивная логика. М., 1978.
17. *Бирюков Б. В., Растригин Л. А., Казаневская В. В., Верстин И. С.* Случайность, случайный поиск и логика // Информационные материалы. Кибернетика. 5 (120). М., 1982. С. 3—51.
18. *Hacking I.* The emergence of probability. Cambridge, 1975.
19. *David F. N.* Games, gods and gambling. L., 1962.
20. *Кац М.* Статистическая независимость в теории вероятностей, анализе и теории чисел. М., 1963.
21. *Мизес Р.* Вероятность и статистика. М.; Л., 1930.
22. *Майстров Л. Е.* Развитие понятия вероятности. М., 1980.
23. *Борель Э., Дельтейль Р., Юрон Р.* Вероятности, ошибки. М., 1972.
24. Принцип симметрии. Историко-методологические проблемы. М., 1978.
25. *Курно О.* Основы теории шансов и вероятностей. М., 1970.
26. *Кац М., Улам С.* Математика и логика. Ретроспектива и перспективы. М., 1971.
27. *Чайковский Ю. В.* Изумительная асимметрия // Знание — сила. 1981. № 2.
28. *Чайковский Ю. В.* Экстремальность как междисциплинарная эвристика // Взаимодействие наук как фактор их развития. Новосибирск, 1988.
29. *Чайковский Ю. В.* Выживание мутантного клона // Генетика. 1977. № 8. С. 1478.
30. *Рота П.* История старинных банковых учреждений. Спб., 1877.
31. *Buffon G. L.* Histoire naturelle. V. 2. P., 1749.
32. *Graunt J.* Natural and political observations made upon the bills of mortality (1662). Baltimore, 1939.
33. *Cardano H.* Opera omnia. T. 1. Lugduni, 1663. P. 262—276.
34. *Adams W. J.* The life and times of the central limit theorem. N. Y., 1974.
35. *Эфрос А. Л.* Физика и геометрия беспорядка. М., 1982.
36. *Уиттл П.* Вероятность. М., 1982.
37. *Тарасов Л. В.* Мир, построенный на вероятности. М., 1984.
38. *Sheynin O. B.* Daniel Bernulli on the normal law // Biometrika. 1970. V. 57. № 1. P. 199—202.
39. *Заславский Г. М.* Стохастичность динамических систем. М., 1984.
40. *Борель Э.* Случай. М.; П., 1923.

THE HISTORY OF SCIENCE AND TEACHING OF SCIENCE (TO THE ANALYSIS OF THE CONCEPTS «CHANCE» AND «PROBABILITY»)

Yu. V. CHAIKOVSKY

The meaning of concepts «chance» and «probability» and the complications concerned with their definition in the context of teaching are discussed. It is indicated, that the circle of problems on the theory of probability is limited and that just with formulation of these limitations the teaching must start. The role of the history of science in the formulation of comprehensive outline of the knowledge is shown by concrete examples.