ДВАДЦАТЬ ВОСЬМАЯ

ХРЕСТОМАТИЯ

По истории теории вероятностей и статистики
составитель О. Б. Шейнин

Берлин
2020

Содержание

oscar.sheynin@gmail.com

Предварительные замечания

Во всём сборнике принято обозначение

S, G, i: скачиваемый документ i на русском языке есть на нашем сайте www.sheynin.de. Сайт перепечатывает Google, см. Oscar Sheynin.

Мы сокращаем давно уже устаревший термин *математическое ожидание* и пишем просто *ожидание*. Этот устаревший термин ввёл Лаплас, чтобы отличить классическое ожидание от ставшего модным (но с тех пор забытого) морального ожидания.

[i] По содержанию брошюры ясно, что читатель должен был быть знаком с анализом в объёме технического вуза, но мог не знать ничего о теории вероятностей (в § 2.4 автор объясняет понятие закона распределения). Мы поэтому сомневаемся, что такой читатель сможет овладеть изложением этой теории сразу на современном языке.

К тому же эта брошюра, как и многие другие научно-популярные сочинения, обращает слишком много внимания на *технологию* в ущерб принципам. Автор лишь косвенно определил устойчивость (Прим. 4.9), ничего не сказал о классическом определении вероятности (которое он назвал классической вероятностью, см. Прим. 2.1), не упомянул хаотического движения (ср. Прим. 1.1), слишком мало сообщил о частотной теории Мизеса и неверно оценил результаты Пуассона по обработке судебной статистики.

Мы не можем упрекнуть автора за скверный стиль, потому что таков недостаток большинства отечественных сочинений (по крайней мере журнальных), но он пронумеровал слишком мало формул, а потому допустил ненужные повторения. Эти недостатки мы частично выправили. В последующих брошюрах стиль автора не улучшился.

[ii] На каких читателей была рассчитана эта брошюра? В трёхтомном *Курсе дифференциального и интегрального исчисления* Г. М. Фихтенгольца, в последнем томе, несколько раз бегло упоминаются интегральные уравнения, но не их собственные функции или ядра, которые появились здесь, в § 3.3. Появились и обобщённые функции (см. Прим. 3.8), и можно было бы добавить сравнительно безвредно упомянутый без пояснения белый шум (конец § 1.4). Читатели, стало быть, должны были быть математиками, и всё же изложение большой части § 3 было слишком беглым даже для них.

Наши примечания показывают, что автор не был достаточно знаком с историей теории вероятностей, а в § 1.1 он повторил своё непростительное заблуждение о пуассоновом исследовании судебной статистики, см.[i, Прим. 3.2].

[iii] Автор был недостаточно знаком с историей теории вероятностей и статистики, а кроме того впряг в свою телегу и коня, и трепетную лань, т. е. почти одновременно описывал и принципы своей науки, и технологию её применения. Лучше было бы, видимо, решительно разделить эти задачи. В Примечаниях мы указали на небрежности изложения и на отсутствие многих необходимых разъяснений, а о стиле автора мы уже упомянули в своих прежних комментариях.

I

В. Н. Тутубалин

Теория вероятностей в естествознании

М., 1972

Введение

Уже со времён Лапласа, Гаусса и Пуассона математическая теория вероятностей пользуется сложным математическим аппаратом[0.1]. В настоящее время в этой науке применяется практически весь классический математический анализ, включая теорию уравнений с частными производными, а со времени классической работы А. Н. Колмогорова [1] также аппарат теории меры и функционального анализа. Тем не менее, книги по теории вероятностей, рассчитанные на массового читателя, обычно начинаются с утверждения, что принципиальные вопросы применения теории вероятностей вполне доступны пониманию неспециалиста. Так считал в 1843 г. Огюстен Курно [2] и с такого же утверждения мы хотим начать эту брошюру.

Можно было бы сказать, впрочем, что эти вопросы недоступны также и для специалиста, поскольку здесь до нашего времени нет полной научной ясности. Точнее говоря, при обсуждении принципиальных вопросов теории вероятностей специалист, вполне владеющий её математическим аппаратом, не имеет никаких преимуществ перед неспециалистом, поскольку математический аппарат здесь не помогает. Здесь важен опыт конкретных применений, приобрести который математику не легче (если не труднее), чем инженеру или исследователю, занятому непосредственными приложениями.

В настоящее время в науке сложились представления об области применимости теории вероятностей, несколько более совершенные, чем во времена Лапласа и Курно, хотя и не столь уж далеко от них ушедшие. Мы начнём с изложения этих представлений.

Глава 1. Всякое ли событие имеет вероятность

1.1 Понятие статистической устойчивости (статистического ансамбля). В учебниках по теории вероятностей, особенно в старых, обычно говорится, что вероятность имеет любое случайное событие, которое может как произойти, так и не произойти. Перечисляется несколько примеров случайных событий, как, скажем, выпадение герба при броске монеты, успешная сдача экзамена студентом, выпадение дождя сегодня вечером. В результате у читателя такого учебника создаётся впечатление, что из того, что мы не знаем, произойдёт данное событие или нет, вытекает возможность говорить о его вероятности.

Таким образом, теория вероятностей оказывается наукой наук, или, по крайней мере, совершенно особой наукой, в которой из полного незнания можно сделать некие содержательные выводы. Естественно, что современная наука энергично отвергает такую трактовку понятия вероятности. Вообще, наука предпочитает эксперименты, результат которых устойчив, т. е. такие, в результате которых всегда наступает (или не наступает) интересующее нас событие. Однако, такой полной устойчивости результатов эксперимента не всегда можно добиться. Например, согласно принятым сейчас в физике взглядам, на квантово-механическом уровне полная устойчивость невозможна[1.1]. Напротив, можно считать достаточно твёрдо установленным, что аккуратный и добросовестный исследователь может во многих случаях добиться если не полной, то статистической устойчивости результатов эксперимента. События, связанные со статистически устойчивыми экспериментами, и составляют, по современным взглядам, область применимости теории вероятностей. Итак, возможность применять теорию вероятностей, вообще говоря, даром не даётся, она служит наградой за большую и кропотливую техническую и теоретическую работу по стабилизации условий, а тем самым и результатов эксперимента.

Что же такое статистическая устойчивость, к которой, как мы только что сказали, следует стремиться? Как определить, удалось ли нам уже достигнуть в своих экспериментах этого желаемого состояния, или нужно что-то ещё совершенствовать?

Следует признать, что в настоящее время на такие вопросы нет исчерпывающего ответа. Некоторые требования статистической устойчивости сформулировал Мизес [3]. Они состоят в следующем. Обозначим через μ_A число наступлений события A при n повторениях некоторого эксперимента. Тогда отношение μ_A/n называется частотой события A. Первое требование статистической устойчивости состоит в том, что при увеличении числа опытов n частота события A становится близкой к некоторому числу $P(A)$, которое называется вероятностью события A. Мизес выражал это в форме

$$\lim \frac{\mu_A}{n} = P(A),\ n \to \infty,$$

но в такой форме это утверждение не может быть проверено экспериментально (поскольку практически нельзя заставить n стремиться к бесконечности).

Второе требование состоит в том, что если заранее, до проведения n опытов, мы уговоримся, что рассмотрим не все опыты, а лишь часть из них (например, опыты с чётными номерами), то частота события A, вычисленная по части опытов, должна быть близкой к тому же самому числу $P(A)$. Конечно, предполагается, что опытов достаточно много.

Скажем сначала о достоинствах формулировки Мизеса. Они состоят, собственно говоря, в том, что её знание позволяет исключить некоторые случаи, когда применение теории вероятностей давало бы ошибочный результат. В этом смысле особенно типично второе требование (первое требование, видимо, хорошо сознают все, кто применяет теорию вероятностей, и здесь ошибок не бывает). Рассмотрим, например, следующий вопрос: можно ли говорить о вероятности того, что изделие, выпущенное данным заводом, окажется бракованным[1.2]. Одной из причин выпуска брака может оказаться не вполне удовлетворительное физическое состояние части рабочих, особенно после праздников.

Следуя второму требованию Мизеса, мы должны сравнить частоту выпуска брака по понедельникам и прочим дням недели, и если эти частоты окажутся заметно различными, то говорить о вероятности выпуска брака не имеет смысла. То же самое относится к выпуску брака в конце квартала или года по сравнению с другими периодами времени (штурмовщина). Наконец, то же самое относится к связи выпуска брака с возможным плохим качеством сырья, с технологическими отклонениями и т. п.

Итак, почти ничего не зная о теории вероятностей, и используя лишь правила Мизеса, мы уже видим, что для её применения к анализу качества продукции необходимо сначала создать хорошо налаженное производство, в котором не было бы ни пьянства, ни прогулов, ни штурмовщины, ни негодного сырья, ни изношенного технологического оборудования, ни пр. Теория вероятностей есть нечто вроде масла к каше: сначала надо иметь кашу.

Впрочем, следует немедленно заметить, что часто наибольшую пользу теория вероятностей приносит не тогда, когда её можно применять, а тогда, когда при попытке её применить обнаруживается отсутствие статистической однородности (*устойчивость* и *однородность* здесь означает одно и то же)[1.3].

Если на продукцию данного завода можно смотреть как на статистически однородную совокупность, то ещё большой вопрос: можно ли улучшить качество продукции (без принципиальных технологических усовершенствований)? Если же качество продукции колеблется (а это нужно устанавливать вероятностными методами), то без сомнения можно установить причину таких колебаний, что возможно улучшит качество.

Главнейшим недостатком приведённой выше формулировки Мизеса является её неопределённость. Не говорится, как велико должно быть число опытов n, чтобы обеспечить заданную близость μ_A/n к $P(A)$. На этот вопрос можно ответить вполне удовлетворительно (что мы и сделаем ниже) лишь при дополнительном предположении независимости результатов отдельных опытов. Экспериментальная проверка независимости отчасти возможна, но трудна и всегда (без исключения) неполна. Но дело обстоит ещё гораздо хуже со вторым требованием Мизеса. В высказанной форме оно попросту противоречиво.

Указывая заранее какую-то часть среди n опытов, мы могли бы случайно напасть на те, в которых событие A наступило (или, напротив, не наступило). Но тогда частота события по части опытов было бы весьма далёкой от его частоты по совокупности всех опытов. Конечно, Мизес имел в виду не указание любой части среди всех опытов, а, скорее, задание разумного правила, выделяющего эту часть. Задание такого правила должно зависеть от наших представлений о том, каким образом может нарушаться статистическая однородность. Например, если мы опасаемся последствий воскресной выпивки, то нужно выделить часть продукции, изготовленную по понедельникам. Если мы желаем проверить независимость события *A* от некоторого другого события *B*, то мы выделяем среди наших опытов две части: опыты, при которых событие B произошло, и не произошло. Но эти разумные соображения плохо переносятся на общий случай, т. е. плохо поддаются формализации в рамках математической теории.

Мы видим, что не существует математически строгого общего способа, чтобы решить, можно ли говорить о вероятности данного события или нельзя. Это, конечно, не означает, что в частном случае мы не можем иметь полной уверенности в применимости вероятностных методов. Например, нет ни малейшего сомнения в возможности вероятностного описания броуновского движения. Но здесь наша уверенность базируется не столько на экспериментальной проверке статистической устойчивости, сколько на общих молекулярно-кинетических представлениях.

В других случаях, как при бросании монеты, наша уверенность основана на опыте бесчисленных игроков в орла и решку. Заметим, впрочем, что вероятность выпадения герба равна 1/2 не казалось очевидным многим выдающимся учёным. Например, Мизес заявлял, что до опыта (априорно) мы этого вообще не знали. Пирсон описал опыт Уелдона, который подбросил 12 костей 26 306 раз, т. е. не считал указанную вероятность очевидной.

Однако, не существует универсального способа для проверки статистической устойчивости (или, как говорят физики, проверки наличия статистического ансамбля). Итак, вероятностный подход никогда не является математически строгим, но достаточно строг физически (поскольку есть статистический ансамбль), и во всяком случае не является менее строгим, чем любое другое применение математических методов в естествознании. Чтобы в этом убедиться, достаточно прочесть § 1, *Что такое энергия*? из четвёртой главы Фейнмана [4, с. 71 – 73]. В прекрасной литературной форме (к сожалению, слишком длинно для цитирования) там сказано, что, хотя в каждом конкретном случае возможно подтвердить закон сохранения энергии (разобравшись, куда она девается), но в современной физике нет общего понятия энергии. Это не мешает нам быть уверенным в законе сохранения энергии настолько, что когда знакомый говорит нам, что в газете появилась статья, где утверждается, что коэффициент полезного действия оказался больше 100%, мы поднимаем на смех и газету, и знакомого.

Многие выводы, полученные приложением методов теории вероятностей к тем или иным статистическим ансамблям, не менее достоверны, чем закон сохранения энергии. Но дело обстоит совсем иначе, когда теория вероятностей применяется, если статистического ансамбля либо явно нет, либо его существование сомнительно. Вообще говоря, в таких случаях современная наука отвергает возможность применения теории вероятностей. Искушение, однако, часто бывает сильным … Рассмотрим, почему это происходит.

1.2. Ограничительность понятия статистического ансамбля (статистической однородности). Дело в том, что концепция статистического ансамбля довольно ограничительна. Рассмотрим упомянутые выше примеры (бросание монеты, сдача экзамена, выпадение дождя). Наличие ансамбля несомненно только в опытах с бросанием монеты. Много хуже обстоит дело в двух других примерах. Можно говорить о вероятности сдачи экзамена наудачу выбранным студентом (ещё лучше, если наудачу выбранный студент в наудачу выбранном институте сдаёт экзамен по наудачу выбранной специальности наудачу выбранному преподавателю). Здесь *наудачу выбранный* означает *выбранный путём эксперимента из статистического ансамбля экспериментов*. Но нельзя говорить о вероятности успешной сдачи данного экзамена данным студентом, так как здесь ансамбль экспериментов состоит ровно из одного невоспроизводимого эксперимента.

Можно говорить о вероятности выпадения дождя в данный день, скажем, 11 мая наудачу выбранного года, но нельзя говорить о вероятности того, что сегодня вечером будет дождь. Дело в том, что если мы утром решаем вопрос о том, будет ли вечером дождь, то обязаны принять в расчёт всю синоптическую обстановку. Другого дня, с такой же синоптической обстановкой (т. е., например, с такой же в точности синоптической картой) мы наверняка не найдём, по крайней мере за то время, в течение которого человечество ведёт метеорологические наблюдения.

В недавнее время появилось много работ по приложениям понятия случайного процесса. Это понятие должно описывать ансамбли таких экспериментов, исходом которых является не событие, даже не измерение (т. е. не единственное число), а функция, например, траектория броуновского движения. Мы не будем говорить о границах применимости понятия случайного процесса, даже если наличие статистического ансамбля бесспорно, мы остановимся на том, что можно сказать в случае, когда нет ансамбля.

Разберём два конкретных вопроса. Первый – из области экономики. Пусть мы наблюдаем за ряд лет (месяцев, недель) значения некоторого экономического показателя, скажем, производительности труда, и хотим сделать прогноз на будущее. Соблазнительно применить для этой цели теорию прогноза случайных процессов. Однако, эксперимент, давший нам наблюдённые значения, очевидно, принципиально невоспроизводим, и никакого статистического ансамбля нет.

Другой вопрос – из области геологии. Допустим, что мы измерили содержание полезного компонента в нескольких точках опробования и хотим определить среднее содержание (а тем самым и запасы, если считать, что геометрические контуры месторождения нам известны). Соблазнительно применить для этой цели теорию оценки среднего случайного процесса. Но и в этом случае неясно, каков должен быть ансамбль реализаций. Если под новой реализацией понимать значения концентрации на другом профиле опробования, то неясно, будут ли они обладать теми же статистическими свойствами. Тем менее ясно, если в качестве других реализаций брать данные по другим месторождениям.

Приведённые примеры достаточно важны, чтобы понять желание создать такие вероятностные методы, которые не нуждались бы в ансамбле. Однако, в современной теории вероятностей таких методов нет (есть только частные методы для спасения концепции статистической однородности, но они далеко не всегда применимы). Как же относятся к применению теории вероятностей в подобных случаях?

1.3. Соотношение между медициной и магией. Поставленный вопрос напоминает вопрос о соотношении между медициной и магией, идею которого мы заимствовали у Фейнмана, но мы разовьём эту тему подробнее. Допустим, что речь идёт о лечении малярии, и шаман знает, что кора хинного дерева помогает, а трясти змеёй над головой больного бесполезно. Тогда он назначает по существу то же лечение, что и врач. Однако, врач вместо коры хинного дерева даёт хинин (это не особенно важно), и, главное, имея представление о жизненном цикле малярийного паразита, правильно назначит продолжительность лечения. Таким образом, врач имеет больше шансов на успех. Но не здесь главное различие между медициной и магией. Главное различие состоит в поведении врача и шамана в случае неудачного лечения. Шаман объяснит неудачу вмешательством дьявола и ничего более эффективного не предпримет, врач же будет искать причину неудачи в надежде, что это поможет, если не данному больному (который, допустим, умер), то хотя бы другим. История науки есть история всё более точного познания действительности, что и ограничивает произвол *дьявола*, перед которым шаман чувствует себя бессильным.

Однако, *дьявола* не удаётся изгнать полностью. Даже в математике возможно его вмешательство. Оно проявляется, например, в противоречиях, и наиболее беспокоят парадоксы теории множеств. В первой половине XX в. была предпринята грандиозная попытка изгнания *дьявола* из математики, связанная с именами Бертрана Рассела, Гильберта, Гёделя и других первоклассных математиков. И что же? Оказалось, что вместе с *дьяволом* приходится изгонять ряд вещей, которых мы никак не хотели бы лишиться, например, представления о числовом континууме. Нельзя, например, утверждать (не протягивая дьяволу палец, вместе с которым он откусит руку), что непрерывная на отрезке функция достигает своего наибольшего значения. Естественно, что такое слишком радикальное изгнание *дьявола* (конструктивный математический анализ) не нашёл признания. *Дьявола* приходится терпеть.

Правда, в математической теории вероятностей этот *дьявол* теоретико-множественных парадоксов есть, собственно, не *дьявол*, а маленький безобидный чертёнок, от которого особого вреда не замечено. Впрочем, автор вспоминает, что когда он однажды хотел применить в доказательстве теоремы трансфинитную индукцию (математический приём, в котором есть нечто от *дьявола*)[1.4], то почувствовал немалое облегчение, увидав, что процесс индукции на самом деле не требует использования трансфинитных чисел, т. е. сводится к обычной математической индукции.

В прикладной теории вероятностей безобидный чертёнок превращается в бойкого рогатого чёрта, излюбленная пакость которого – нарушать статистическую однородность. Пока мы держимся понятия ансамбля и проверяем статистическую однородность доступными нам средствами, мы в состоянии хотя бы своевременно заметить дьявольскую пакость. Оставив это понятие, мы полностью отдаёмся во власть *дьявола* и должны быть готовы к неожиданностям. Таким образом, с точки зрения современной теории вероятностей граница между наукой и магией определяется понятием статистического ансамбля. Следовательно, выводы, полученные путём применения теории вероятностей при отсутствии статистического ансамбля экспериментов, не обладают научной достоверностью.

Аппарат науки, в отличие от аппарата магии, должен быть состоятелен во всех своих частях. Между тем, когда делают вывод, например, что ошибка результата, полученного по единственной реализации случайного процесса, находится в заданном интервале с вероятностью 0,95, то к чему относится эта вероятность? К ансамблю реализаций, который мы должны примыслить к единственной наблюдаемой реализации, чтобы воспользоваться понятием случайного процесса? Но нам нет дела до всех прочих реализаций. Вообще говоря, ничего не стоит привести примеры неверных выводов, полученных путём применения теории вероятностей и в экономике, и в геологии, да и в других науках, где бессмысленно говорить о статическом ансамбле.

Исторически наука выросла из магии, но относится к ней с пренебрежением и желала бы её игнорировать. Нельзя, однако, полностью поддаться и такому искушению. С точки зрения представителя конструктивного направления в математике, магией является обычный математический анализ. Следует скорее различать между белой и чёрной магией. Последняя связана с субъективной недобросовестностью. Добросовестные исследования, в которых пытаются прилагать теорию вероятностей при отсутствии статистического ансамбля, в настоящее время не могут игнорироваться. Можно рискнуть сделать прогноз, что нечто из того, что сегодня является магией, завтра станет наукой. Было бы неразумно слишком крепко держаться за установившуюся концепцию статистической однородности. Однако, мы здесь будем полностью придерживаться указанной концепции, поскольку в настоящее время нет иного способа получения действительно надёжных результатов.

1.4. Резюме. Итак, в процессе совершенствования экспериментов в той или иной отрасли науки (техники) часто возникает особое положение, когда полной устойчивости результатов эксперимента добиться не удаётся, но возникает явление статистической устойчивости. Она характеризуется устойчивостью частот наступления различных событий, связанных с исходом эксперимента. Исчерпывающая проверка устойчивости частот (синонимы: статистической однородности, статистического ансамбля) невозможна. Однако, во многих случаях наличие статистического ансамбля достаточно достоверно. Эти случаи и составляют, по современным взглядам, область научного применения теории вероятностей.

Тем не менее, имеется хорошо понятное желание применять теорию вероятностей и в случаях, когда результаты экспериментов имеют неопределённость, но нельзя говорить о статистически однородном ансамбле. Такие применения пока что относятся не к науке, а к магии, однако при субъективной добросовестности их нельзя игнорировать. Возможно, что в будущем их удастся сделать научными. Как указывает вся история науки, её возникновение происходит на основе фактического материала, накопленного при занятиях магией.

Глава 2. Основы математического аппарата теории вероятностей

Современная теория вероятностей резко делится на математическую и прикладную. К первой примыкает математическая статистика, вторая тесно связана с так называемой *прикладной статистикой*. Попытка дать *определения* указанных наук завела бы нас в такие схоластические дебри, что мы с ужасом отказываемся от неё. Мы хотим занять некоторую промежуточную позицию и начнём с математической теории вероятностей.

Это направление изучает следствия из аксиоматики Колмогорова [1] и далеко продвинулось в разработке чисто математических методов. Но оно совершенно оставляет в стороне вопрос, каким явлениям действительности аксиоматическая модель соответствует хорошо, каким похуже, а каким и вовсе не соответствует. Можно привести поистине анекдотические примеры ошибок, в которые впадали математики, не имеющие достаточного опыта и практической интуиции при попытке работать в приложениях.

Тем не менее, для развития математического аппарата аксиоматическая модель удобна. Известные понятия и теоремы теории вероятностей в её рамках превращаются просто в частные случаи соответствующих понятий и теорем анализа. В настоящей главе мы изложим как раз этот материал, а содержательным теоремам теории вероятностей посвящены следующие главы.

Данная книга – не учебник. Поэтому изложение теории вероятностей здесь кратко и местами конспективно. Хотя формально знание теории вероятностей не предполагается и все понятия, нужные для понимания дальнейших глав книги, определяются, но нашему изложению не хватает достаточного числа примеров. Без них нельзя научиться прилагать аксиоматическую модель. Таким образом, будет лучше, если читатель уже познакомился или собирается познакомиться с теорией вероятностей по любому учебнику, хотя бы и не придерживающемуся аксиоматической трактовки. Из современных учебников мы особенно рекомендуем часть 1 книги Феллера [5].

2.1. Дискретное пространство элементарных событий. В простейшем случае, который вполне достаточен для решения многих задач, вся теория вероятностей состоит из одного понятия, одной аксиомы и одного определения. Мы перечислим их.

Понятие вероятностного пространства. Вероятностным пространством называется любое конечное или счётное множество, элементам которого поставлены в соответствие неотрицательные числа, называемые их вероятностями. Слово *множество* означает то же, что *совокупность*, т. е. нечто, составленное из отдельных элементов. Множество называется счётным, если его элементы можно занумеровать числами 1, 2, …, n, …

Нам необходимо ввести соответствующие обозначения. Вероятностное пространство мы будем обозначать буквой Ω; его элементы $\omega_1, \omega_2, ..., \omega_n, ...$ Будем записывать это одним из следующих способов:

$$\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}, \text{ или } \Omega = \{\omega_i : i = 1, 2, ..., n, ...\}.$$

Вероятности элементов обозначим через $P(\omega_i) \geq 0$. Элементы множества Ω мы будем называть также элементарными событиями или элементарными исходами.

Аксиома. Сумма вероятностей всех элементарных событий равна единице, или, во введённых обозначениях,

$$P(\omega_1) + ... + P(\omega_n) + ... = \sum_{i=1}^{\infty} P(\omega_i) = \sum_{\omega_i \subseteq \Omega} P(\omega_i) = 1.$$

Определение. Событием называется любое подмножество множества элементарных событий; вероятностью события называется сумма вероятностей элементарных событий, входящих в это событие.

В математических обозначениях тот факт, что множество A является подмножеством множества Ω (т. е., что A составлено из некоторых элементов, входящих в Ω) записывается в виде $A \subseteq \Omega$. Вероятность события A обозначается через $P(A)$, и данное определение записывается в виде

$$P(A) = \sum_{\omega_i \subseteq A} P(\omega_i).$$

Запись под знаком суммы означает, что нужно суммировать $P(\omega_i)$ по тем и только тем ω_i, которые входят в A.

Изложенную математическую модель можно применить для решения очень многих задач по теории вероятностей. Однако, все задачи этой теории первоначально формулируются не в терминах пространства элементарных событий, т. е. не на аксиоматическом, а на обычном языке. Это неизбежно, так как только по задачам знакомится всякий изучающий теорию вероятностей с теми конкретными ситуациями, в которых её можно применять. На аксиоматическом языке описывать конкретные ситуации невозможно. Итак, необходимо уметь переводить условие задачи на язык пространства элементарных событий.

Положение здесь вполне аналогично тому, с которым школьник сталкивается при решении задач на составление систем уравнений: он тоже должен перевести условие задачи с одного языка на другой. Такой перевод может быть совсем лёгким, но может быть и трудным. Возможна и неоднозначность перевода: одной и той же задаче могут соответствовать различные системы уравнений. Может получиться и так, что одна система трудно составляется, но легко решается, другая же – наоборот. С аналогичными явлениями мы сталкиваемся и в задачах по теории вероятностей.

Итак, подчеркнём, что при введении пространства элементарных событий, отвечающего данной конкретной задаче, речь идёт не о некоторой чисто математической операции, вроде доказательства теоремы, а именно о переводе с одного языка на другой. Поэтому здесь бессмысленно добиваться той строгости, которая принята в математике. Чёткие математические формулировки исчерпаны, идут правила перевода.

В задачах по теории вероятностей речь обычно идёт о некотором эксперименте, и множество Ω состоит из всех возможных исходов эксперимента. Например, при бросании монеты Ω состоит из двух элементарных исходов

$$\Omega = \{\text{герб; решётка}\},$$

а при броске игральной кости – из шести таких исходов

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Подбросим две кости, и Ω будет состоять из всех пар вида

(m, n), т. е. из сумм очков на каждой кости:

$$\Omega = [(m, n): m = 1, ..., 6; n = 1, ..., 6].$$

Примером события A здесь может оказаться чётная сумма очков

$$A = [(m, n): m = 1, ..., 6; n = 1, ..., 6; m + n \text{ чётно}\}.$$

В общем, описание множества элементарных событий обычно нетрудно, но совсем другое дело, определение по условиям задачи вероятностей $P(\omega_i)$ отдельных элементарных событий. В соответствии с частотной концепцией вероятности для определения $P(\omega_i)$ следовало бы много раз произвести эксперименты и в качестве приближённых значений этих $P(\omega_i)$ принять частоты наступления элементарных исходов ω_i. Это не всегда возможно и на самом же деле определение вероятностей по частотам является сложной задачей. Поэтому большую роль играют те случаи, когда вероятности можно определить без экспериментов, по каким-то умозрительным соображениям. Например, довольно часто множество Ω содержит конечное число N элементов, вероятности которых с несомненностью представляются равными. В соответствии с аксиомой вероятность каждого элементарного события в этом случае равна $1/N$, а если A содержит M элементарных событий, то

$$P(A) = \sum_{\omega_i \subseteq A} P(\omega_i) = \frac{M}{N}. \qquad (2.1.1)$$

Формулу (2.1.1) выражают словами так: вероятность события равна отношению числа благоприятных исходов (т. е. исходов, входящих в событие), к числу всех возможных. Если формула (2.1.1) применима, то говорят, что рассматриваемая задача это задача на классическую вероятность[2.1].

Итак, в современной трактовке формула (2.1.1) не является определением вероятности, она применима лишь в частном случае, когда все элементарные события равновероятны. Но в каких же случаях они равновероятны? Это достаточно деликатный вопрос. Например, продолжительные эксперименты с игральными костями показывают, что различные грани кости, вообще говоря, не равновероятны (трудно изготовить совершенно симметричную кость). С другой стороны, специальные меры, принимаемые при розыгрышах лотерей, делают одинаковыми вероятности выигрыша различных билетов.

Для иллюстрации возможностей математической модели мы рассмотрим задачу о жеребьёвке, предполагая, что были приняты достаточные меры, чтобы можно было пользоваться понятием классической вероятности.

Задача о жеребьёвке. При распределении квартир в жилищных кооперативах[2.2] жеребьёвка иногда производится в два этапа. Сначала разыгрывается порядок, в котором члены кооператива будут подходить к урне, содержащей билетики с номерами квартир, а затем извлекаются сами билетики. Нужны ли эти два этапа? Иначе говоря, кто имеет больше шансов вытащить более подходящую квартиру, – тот, кто подходит за билетиками первый, или тот, кто оказывается последним?

Решение. Пусть имеется N квартир, из которых квартиры с номерами 1, 2, .., n плохие, а квартиры с номерами $n + 1, ..., N$, хорошие. Найдём вероятность того, что член кооператива, подошедший к урне k-м по порядку, вытащит плохую квартиру. Опыт, состоящий в проведении жеребьёвки, т. е. в вынимании N билетиков из урны, очевидно, имеет исходы

$i_1, i_2, ..., i_N$, все i_k различны.

Здесь i_1 это номер квартиры, вынутый первым, подошедшим к урне, i_2, это номер, вынутый вторым и т. д. Для i_1 имеется N разных возможностей, для i_2, $(N – 1)$ возможностей, поскольку $i_2 \neq i_1$, для i_3, $(N – 2)$, поскольку $i_3 \neq i_1$ и $i_3 \neq i_2$ и т. д. Всего элементарных событий имеется

$$N(N-1)...2\cdot 1 = N!.$$

Если билетики в урне перемешаны тщательно, то все элементарные события должны быть равновероятны, и мы имеем задачу на классическую вероятность. Пусть теперь A_k будет событием, состоящим в том, что k-й член кооператива вытащил плохую квартиру. Иначе говоря, A_k состоит из таких элементарных событий $i_1, i_2, ..., i_N$, что i_k принимает одно из значений 1, 2, ..., n при любых $i_1, ..., i_{k-1}, i_{k+1}, ..., i_N$. Подсчитаем число таких элементарных событий. Имеем:

для i_k имеется n возможностей,
для $i_1, i_2, ..., i_{k-1}$ имеется $N – 1, N – 2, ..., N – k + 1$ возможностей,
для $i_{k+1}, ..., i_N$ имеется $(N – k), ..., N – (N – 1) = 1$ возможностей.

Перемножая все возможности, получаем, что событие A_k состоит из

$(N – 1)/n = (N!/N)n$ элементарных событий, поэтому

$$P(A_k) = \frac{(N!/N)n}{N!} = \frac{n}{N}$$

и не зависит от k.

Итак, вероятность вытащить плохую квартиру не зависит от места в очереди, и первая жеребьёвка, таким образом, излишня. Но наш вывод получен в предположении, что билетики в урне перемешаны достаточно тщательно. Если это не так, то шансы членов кооператива при различных извлечениях не равны, и первая жеребьёвка значительно выравнивает шансы. К сожалению, неизвестно, как именно надо перемешивать билетики, чтобы обеспечивать равенство шансов. Способ, принятый при розыгрыше лотерей, слишком хлопотлив и поэтому нельзя признать двухэтапную жеребьёвку совершенно излишней[2.3].

2.2. Условная вероятность. Читатель, знакомый с урновыми моделями теории вероятностей, несомненно заметил, что модель пространства элементарных событий вполне изоморфна модели с извлечением шаров из урны, с той только разницей, что у нас разные элементарные события могут иметь различные вероятности, а число этих событий может быть и бесконечным. Действительно, настоящая роль аксиоматики Колмогорова выясняется лишь при рассмотрении более, чем счётных пространств элементарных событий. Но даже в простейших (конечном или счётном) случаях преимущество аксиоматического подхода состоит в том, что он чётко разделяет на две части решение задач из теории вероятностей: **1.** Выбор математической модели явления или эксперимента. **2.** Вычисления в рамках этой модели.

Таким образом, мы здесь следуем совету Декарта, делим каждую стоящую перед нами задачу на столько частей, чтобы частичные задачи можно было решить.

Первая часть, выбор математической модели, бесспорно является труднейшей. Как уже говорилось, трудность состоит в определении вероятностей элементарных событий. Для формулировки более или менее общих рецептов, чтобы преодолеть эту трудность, вводятся некоторые новые понятия, и мы уже рассмотрели понятие классической вероятности. Другим полезным понятием является понятие условной вероятности. Но сначала целесообразно рассмотреть обычные операции над событиями. В принятой сейчас теоретико-множественной трактовке эти операции совпадают с операциями теории множеств.

Сумма (*объединение*) *событий*. Суммой $A \cup B$ событий A и B называется событие, состоящее из тех элементарных событий, которые входят в событие A или B (или в то и другое).

Произведение (*пересечение*) *событий*. Произведением AB событий A и B называется событие, состоящее из тех элементарных событий, которые входят в оба события A и B.

Дополнение события (*противоположное событие*). Дополнением $\bar{A}$ события A является событие, состоящее из тех элементарных событий, которые не входят в событие A.

Если опыт кончается одним из элементарных исходов, входящих в некоторое событие C, то говорят, что событие C наступило. Таким образом, сумма событий A и B наступает в том случае, когда наступает хотя бы одно из этих событий. Произведение AB наступает в том случае, когда наступили оба эти события. Дополнение $\bar{A}$ к событию A наступает в том случае, когда A не наступает.

Математически условная вероятность $P(A/B)$ того, что событие A наступает при условии, что событие B наступило, определяется равенством

$$P(A/B) = \frac{P(AB)}{P(B)},\ P(B) \neq 0.$$

Отсюда вытекает, что $P(AB) = P(B)\,P(A/B)$.

Роль понятия условной вероятности раскрывается его частотным истолкованием. Рассмотрим n опытов, в каждом из которых могли наступить или не наступить события A и B. Обозначим через μ_A, μ_B, и μ_{AB} числа наступления событий A, B, и AB. Очевидно, что μ_{AB} есть также число наступлений события A в тех опытах, в которых B наступило, а отношение μ_{AB}/μ_B – условная частота события A при условии, что событие B наступило. Имеем

$$\frac{\mu_{AB}}{\mu_B} = \frac{\mu_{AB}/n}{\mu_B/n} \approx \frac{P(AB)}{P(B)} = P(A/B)$$

и условная вероятность понимается как условная частота.

Пусть пространство элементарных событий Ω разбито на части $B_1, B_2, ..., B_n$, так что $\Omega = B_1 \cup B_2 \cup ... \cup B_n$ и никакие два множества B_i и B_j не имеют общих элементов. Тогда, для любого события $A \subseteq \Omega$ имеем

$$A = AB_1 \cup AB_2 \cup ... \cup AB_n,$$

т. е. элементарные события, входящие в A, разбиваются на входящие в $B_1, B_2, ..., B_n$, и очевидно

$$P(A) = P(AB_1) + P(AB_2) + ... + P(AB_n).$$

Это вытекает из определения $P(A)$, см. § 2.1.

По определению условной вероятности

$$P(A) = \sum_{i=1}^{n} P(AB_i) = \sum_{i=1}^{n} P(B_i)P(A/B_i) \qquad (2.2.1)$$

(формула полной вероятности).

Имеет место и так называемая формула Бейеса

$$P(B_i/A) = \frac{P(AB_i)}{P(A)} = \frac{P(B_i)P(A/B_i)}{\sum_{i=1}^{n} P(B_i)P(A/B_i)}. \qquad (2.2.2)$$

Как формулу полной вероятности, так и формулу Бейеса мы получили из определения условной вероятности путём совершенно элементарных преобразований. Следовательно, они не могут считаться содержательными математическими теоремами. Тем не менее, их роль велика. Рассмотрим сначала применение первой из них. Пусть, для определённости, событие A означает дефектность некоторого изделия. Пусть найдено, что само событие A не обладает статистической устойчивостью, но имеются несколько взаимно исключающих друг друга условий производства $B_1, B_2, \ldots, B_n$ таких, что при фиксированных условиях B_i можно говорить о вероятности $P(A/B_i)$, т. е. что имеется статистическая устойчивость.

Допустим теперь, что продукция, выпущенная при условиях $B_1, B_2, \ldots, B_n$, смешана в общую кучу, причём в ней доля изделий, выпущенных при условиях B_i равна известному числу $P(B_i)$. Рассмотрим опыт, при котором из кучи извлекается наудачу и проверяется одно изделие. Этот опыт имеет два исхода: А (брак) и $\bar{A}$ (годное изделие). Поскольку извлечение производится наудачу, этот опыт будет статистически устойчив. При этом $P(A)$ выражается формулой (2.2.1) и

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

С формулой Бейеса (2.2.2) раньше связывались необоснованные надежды, поскольку не исключалась возможность субъективной трактовки вероятности. Например, если мы занимаемся наукой и имеем несколько научных гипотез $B_1, B_2, \ldots, B_n$, в которые верим с вероятностями $P(B_1), P(B_2), \ldots, P(B_n)$, то для выяснения вопроса, какая из гипотез верна, желательно сделать опыт. Пусть некоторое событие A наступит при этом с вероятностью $P(A/B_i)$, если на самом деле верна гипотеза B_i. Вычисляя по формуле (2.2.2) вероятности $P(B_i/A)$, мы получим новые оценки правдоподобия различных гипотез.

В современной теории вероятностей субъективная вероятность считается понятием магии[2.4], сохранилась лишь терминология, согласно которой вероятности $P(B_1), P(B_2), \ldots, P(B_n)$ называются априорными, а вероятности $P(B_1/A), P(B_2/A), \ldots, P(B_n/A)$, апостериорными. И всё же к магии надо относиться осторожно: существует важная научная область, в которой вышеприведённое магическое рассуждение воскресает в бесспорно научном виде. Речь идёт о машинной диагностике.

Пусть в данную клинику принимаются больные с заболеваниями $B_1, B_2, \ldots, B_n$. Априорные вероятности $P(B_1), P(B_2), \ldots, P(B_n)$ истолковываются как частоты встречаемости соответствующих болезней. Под событием A нужно понимать совокупность результатов обследования больного, проводимого с диагностической целью. Апостериорные вероятности $P(B_i)/A)$ дают некоторый объективный способ суммирования информации, содержащейся в обследованиях A. Объективный не значит обязательно хороший, но заранее им нельзя пренебрегать. Вопрос лишь в том, как узнать вероятности $P(A/B_i)$, нужные для вычисления апостериорных вероятностей.

Казалось бы, для того, чтобы узнать статистическим путём вероятности нужно лишь взять по материалам историй болезни пациентов с болезнью B_i частоту *A*. Но здесь мы сталкиваемся с весьма неприятной неожиданностью. *A* есть результат большого числа обследований, так сказать, совокупность всех существенных для диагноза признаков, которые можно выявить у данного больного. В настоящее время самое простейшее медицинское обследование включает ряд анализов и исследований, а также осмотр многими врачами-специалистами. Никак не будет преувеличением предположить, что объём информации окажется таким, что для записи *A* двоичным кодом потребуется 50 двоичных знаков. Собственно говоря, в 50 двоичных знаков может быть удастся уложиться только после тщательного отбора признаков лишь существенных для диагноза.

Если мы примем эту цифру 50, то получим, что для *A* имеется $2^{50} \approx 10^{15}$ различных возможных значений. Если считать, что на основании прошлой статистики накоплен материал по 10^4 больным, то на каждое возможное значение *A* в среднем придётся $10^4 \text{x} 10^{-15} = 10^{-11}$ наблюдений. Практически это значит, что для подавляющего большинства значений *A* нет никаких наблюдений. Иными словами, почти каждый вновь поступающий в клинику больной будет иметь ранее не встречавшийся результат обследований *A*. Таким образом, вероятности $P(A/B_i)$ совершенно невозможно определить непосредственно по частотам.

Вообще, при практических статистических исследованиях мы обычно заходим в тупик, если желаем рассматривать сразу много факторов и связи между ними. При классификации по нескольким признакам статистический материал очень скоро распадается на группы, содержащие по одному наблюдению, с которым неизвестно, что делать дальше.

Таким образом, формула Бейеса, будучи математической тривиальностью, естественно не может сама по себе дать какого-либо практического эффекта. Тем не менее, учёт многих факторов в медицине возможен (существуют работы, результаты которых трудно подвергнуть сомнению), но об этом ещё рано писать для широкого читателя. В частности, одна из возможностей связана с применением понятия независимости, к формулировке которого мы и переходим.

2.3. Независимость. Когда мы желаем рассматривать события $A_1, A_2, ..., A_n$, то для их полной вероятностной характеристики нужно задать вероятности всевозможных наборов

$$P(C_1C_2, ...C_n),$$

где каждое C_i может принимать два значения, A_i и $\bar{A}_i$. Всего, как нетрудно подсчитать, нужно задать 2^n вероятностей. Это число очень быстро возрастает с ростом *n*, и для задания такого количества вероятностей весьма скоро возможности любого эксперимента становятся недостаточными. На практическое применение могут рассчитывать только такие вероятностные модели, в которых каким-то образом преодолена эта трудность. Основную роль здесь играет понятие независимости.

Определение. Два события, A и B, называются независимыми, если условная вероятность $P(A/B)$ совпадает с безусловной. Иначе говоря,

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

Для n событий $A_1, A_2, ..., A_n$ независимость определяется равенством

$$P(C_1C_2 \ldots C_n) = P(C_1)P(C_2) \ldots P(C_n), \qquad (2.3.1)$$

где каждое C_i может принимать значения A_i и $\overline{A}_i$. Поскольку $P(\overline{A}_i) = 1 - P(A_i)$, для задания вероятностей (2.3.1) в случае независимых событий нужно задать лишь n значений $P(A_1)$, $P(A_2)$, ..., $P(A_n)$.

Независимые события действительно существуют. Они осуществляются в экспериментах, проводимых независимо друг от друга (в обычном физическом смысле). В учебнике по теории вероятностей следовало бы показать читателю, как строится соответствующее пространство элементарных событий. Но эта книга не учебник. Мы достаточно подробно изложили самые основные понятия теории вероятностей, чтобы показать, как коротко и удобно (одно понятие, одна аксиома, одно определение) это делается на теоретико-множественном языке. Дальнейшее развитие тоже делается коротко и удобно, но от стиля учебника мы переходим к стилю конспекта.

2.4. Случайные величины. *Определение*. Случайной величиной называется функция, определённая на множестве элементарных событий.

Обычно случайные величины обозначаются греческими буквами ξ, η, ζ и т. д. Если желательно включить обозначение аргумента $\omega \subset \Omega$, то пишут ξ(ω), η(ω), ζ(ω) и т. д.

С каждой случайной величиной $\xi = \xi(\omega)$ связывается набор её возможных значений $a_1, a_2, ..., a_n, ...$ (все эти числа различны) события вида

$$\{\omega\colon\ \omega \subset \Omega,\ \xi(\omega) = a_i\} = \{\xi = a_i\},$$

т. е. события, состоящие в том, что случайная величина примет одно из своих значений), и вероятности

$$P\{\xi = a_i\} = \sum P(\omega) = p_i,\ \ \omega{:}\xi(\omega) = a_i.$$

Таблица

$a_1 \quad a_2 \quad \ldots\ a_n \ldots$

$p_1 \quad p_2 \quad \dots p_n \dots$

называется распределением случайной величины ξ.

Надо отчётливо представить себе, что практически почти всегда имеют дело не с самими случайными величинами, а только с их распределениями. Причина этого состоит, коротко говоря, в том, что случайные величины как функции элементарных событий обычно не наблюдаемы. В результате эксперимента, который закончился одним из элементарных событий ω, мы обычно узнаём значение случайной величиные ξ(ω), но не узнаём значения ω.

Рассмотрим, например, бросание игральной кости, но введём пространство элементарных событий следующим сложным образом: будем понимать под ω набор значений координат и скоростей кости в тот момент, когда мы её выпускаем из рук. Точнее говоря, будем считать, что ω это набор указанных чисел, записанных с такой точностью, чтобы по ним можно было бы однозначно определить число выпавших очков ξ(ω). В настоящее время в рамках микромира такую однозначную определённость мы не можем признать, но при движении кости мы в этом не сомневаемся. Впрочем, никто этого не проверял, и, во всяком случае, наблюдать ω с такой точностью чрезвычайно трудно, и даже технически (не но принципиально) невозможно. Наблюдать же ξ(ω) легко, и этим только и занимаются игроки в кости.

Итак, пространство элементарных событий Ω чрезвычайно удобно концептуально, как мы уже видели и ещё увидим, но, как правило, фактически не наблюдаемо. События же вида $\{\xi = a_i\}$ наблюдать легче. И, однако, этих событий всё-таки слишком много, поэтому предпочитают характеризовать распределение случайной величины несколькими параметрами, т. е. функциями от значений a_i и вероятностей p_i.

Рассматривают не произвольные распределения, а такие, которые однозначно определяются небольшим числом параметров. Хорошо, если их один-два, терпимо, если три-четыре. Если же кто-либо определяет из опыта более четырёх параметров, то в его результатах сомневаются. Дело в том, что если подобрать слишком много параметров, то, как эмпирически замечено, можно любые данные наблюдений подогнать под любой закон определения.

Важнейшим из параметров распределения является ожидание. Мы дадим его определение не в той форме, как это принято. Обычное определение будет у нас очень простой теоремой.

Определение. Ожиданием случайной величины ξ = ξ(ω) называется число Eξ, определяемое формулой

$$\mathrm{E}\xi = \sum \xi(\omega) P(\omega), \omega \in \Omega.$$

Здесь предполагается, что ряд в правой части абсолютно сходится, т.е., что

$$\mathrm{E}(\xi) = \sum_{\omega_i \subseteq A} |\,\xi(\omega)\,|\, P(\omega) < \infty.$$

Если это не так, то говорят, что эта случайная величина не имеет ожидания. Нетрудно убедиться в том, что наше определение на самом деле совпадает с общепринятым. Действительно, пусть [следуют формулы, которые сложно копировать. Результат:]

$$\mathrm{E}\xi = \sum a_i p_i, \text{ суммирование производится по } a_i. \qquad (2.4.1)$$

Это является обычным определением ожидания.

Но с помощью нашей формы определения удобнее доказывать теоремы о свойствах ожидания. Докажем, например, что для любых двух случайных величин ξ и η имеем

$$\mathrm{E}(\xi + \eta) = \mathrm{E}\xi + \mathrm{E}\eta,$$

предполагая, что оба члена в правой части существуют. [Мы снова опускаем сложные формулы.]

Во многих учебниках только что доказанное нами утверждение доказывается неверно. Ход доказательства при этом таков. […]

Затруднения подобного рода, встречаемые при доказательстве ряда теорем, и заставили нас несколько изменить форму определения ожидания. Однако, формула (2.4.1) чрезвычайно важна: она показывает, что ожидание выражается через распределение случайной величины.

Вторым важнейшим параметром распределения случайной величины является дисперсия.

Определение. Дисперсией Dξ случайной величины ξ называется число, определяемое равенством

$$\mathrm{D}\xi = \mathrm{E}[(\xi - \mathrm{E}\xi)^2].$$

[Автор выводит формулу:]

$$\mathrm{D}\xi = \sum(a_i - \mathrm{E}\xi)^2 p_i, \text{ суммирование происходит по } a_i.$$

Для случайных величин, как и для событий, важнейшую роль играет понятие независимости. Мы дадим определение независимости в совокупности для трёх случайных величин. Для любого числа случайных величин оно вводится аналогично.

Определение. Случайные величины ξ, η, ζ, принимающие соответственно значения

$$a_1, a_2, \ldots;\ b_1, b_2, \ldots;\ c_1, c_2, \ldots,$$

называются независимыми, если

$$P(\xi = a_i, \eta = b_j, \zeta = c_k) = P(\xi = a_i)P(\eta = b_j)P(\zeta = c_k)$$

для любых a_i, b_j, c_k.

Докажем, что для независимых случайных величин имеет место соотношение

$$E(\xi \cdot \eta \cdot \varsigma) = E\xi \cdot E\eta \cdot E\varsigma.$$

[...] Из доказанного утверждения без труда выводится тот факт, что дисперсия суммы независимых случайных величин равна сумме дисперсий слагаемых. Этот вывод имеется в любом учебнике.

Мы закончили изложение основных понятий теории вероятностей для дискретного случая, когда опыт имеет лишь конечное или счётное число элементарных исходов. Нам предстоит сейчас посмотреть, что происходит, если опыт естественнее описывать более сложным пространством.

2.5. Переход к случаю общего пространства элементарных событий. Если результатом опыта является некоторое измерение, то можно заявить, что, поскольку все измерения делаются лишь с конечной точностью, множество элементарных исходов всё равно окажется не более, чем счётным. Однако, история развития науки показывает, что физические теории много выигрывают от рассмотрения непрерывных моделей, в которых результат опыта может быть любым числом. Только в таких моделях можно пользоваться дифференциальными уравнениями. Читатель, знакомый с разностными уравнениями, легко представит себе, насколько дифференциальные уравнения изящнее и проще.

Таким образом, хотя в современной физике и есть некоторые неясные представления о возможной дискретности пространства, несомненно, что от представления о континууме не так-то просто отказаться. Какой же должна быть теория вероятностей, учитывающая представления о континууме? Ответ на этот вопрос даётся знаменитой аксиоматикой Колмогорова, см. [1] или [5]. В её основе лежит представление о пространстве элементарных исходов Ω, которое теперь может быть произвольным. Некоторые (но не все!) подмножества множества Ω объявляются, так сказать, *наблюдаемыми* в результате эксперимента и называются событиями.

Если A событие, то мы можем сказать, наступило оно, или нет в результате эксперимента, и в этом смысле оно наблюдаемо. Следовательно, можно говорить о частоте его наступления, а тогда уж и о вероятности $P(A)$.

Основное требование аксиоматики Колмогорова (в котором содержится, как в зародыше, достоинства и недостатки всей теории) состоит в том, что если у нас имеется счётное множество событий $A_1, A_2, ..., A_n, ...$, то их сумма и пересечение также являются событиями. Кроме того, предполагается, что Ω – событие, причём $P(\Omega) = 1$, и что дополнение $\overline{A}$ к любому событию A также есть событие. Относительно вероятности предполагается следующее фундаментальное свойство: если события $A_1, A_2, ..., A_n, ...$ не пересекаются (не имеют общих элементов), то

$$P[\bigcup_{i=1} A_i] = \sum_{i=1} P(A_i), \qquad (2.5.1)$$

где символ U обозначает сумму. Это утверждение можно в дискретном случае объявить теоремой, выводя его из определения § 2.1. В общем случае это аксиома, а упомянутое определение не действует.

Разберём на конкретном примере, к чему обязывает нас использование колмогоровской аксиоматики. Рассмотрим эксперимент, состоящий в случайном бросании точки на отрезок [0, 1]. В этом случае пространство элементарных исходов Ω, очевидно, должно состоять из всех точек этого отрезка. Если $0 \le a < b \le 1$, то нам было бы крайне досадно, если бы нельзя было говорить о вероятности попадания случайной точки ω внутрь отрезка [*a*, *b*]. Итак, множества вида

$$\{\omega: a \le \omega \le b\}$$

мы хотим считать событиями. Будем считать, что

$$P\{\omega: a \le \omega \le b\} = b - a,$$

т. е., что вероятность попадания случайной точки на отрезок равна длине этого отрезка. Пока всё совершенно естественно.

Но теперь мы должны допустить, что событиями будут не только отрезки, но и всё, что можно получить из них суммированием и пересечением в счётном числе (а также взятием дополнения). Беря точку *c*, $0 \le c \le 1$, и последовательность отрезков вида $[c - 1/n, c + 1/n]$, мы увидим, что пересечение этих отрезков в счётном числе состоит из одной точки *c*. Значит, любая точка – событие. Множество рациональных точек складывается из счётного числа точек, и, следовательно, это множество – событие. Но множество иррациональных точек есть дополнение к множеству рациональных точек, а, следовательно, также есть событие. Итак, мы считаем наблюдаемым, является ли брошенная на отрезок точка рациональной или иррациональной, хотя физически это никак не наблюдаемо. Мы видим, что колмогоровской моделью надо пользоваться осторожно, иначе из неё можно извлечь физически нелепые следствия[2.4].

Особенно сложные варианты такой модели применяются в теории случайных процессов. Здесь от исследователя требуется особая осторожность, наличие, так сказать, определённого естественнонаучного вкуса, так как иначе легко вывести из принятой математической модели такие результаты, которые в лучшем случае не имеют никакого физического истолкования, а в худшем подают повод к неправильной интерпретации. Для примера упомянем о математической теореме, согласно которой коэффициент диффузии броуновского движения может быть найден совершенно точно, если известна траектория на любом сколь угодно малом интервале времени. Можно встретить точку зрения, согласно которой в силу этой теоремы практическая оценка коэффициента диффузии не представляет никаких сложностей. Эта точка зрения даже до некоторой степени утвердилась в литературе по статистике случайных процессов, но она совершенно ошибочна.

Применению указанной теоремы к реальному броуновскому движению мешают два обстоятельства. Во-первых, на малых интервалах времени математическое *броуновское движение*, т. е. винеровский процесс, не описывает реального броуновского движения, а в оценку коэффициента диффузии входят как раз изменения положения *частицы* за бесконечно малые промежутки времени. Во-вторых, совершенно нереалистично представление о том, что траектория какого-то случайного процесса была бы нам точно известна на каком-то интервале времени. Мы никак не знаем способа точно задавать нерегулярно меняющуюся функцию, которая не описывается аналитическим выражением, но мы не имеем возможности останавливаться более подробно на теории случайных процессов.

Вернёмся к вероятности *P*. На отрезках она совпадает с их длиной, однако, из отрезков можно построить весьма сложные множества. Математическая корректность требует, чтобы вероятность *P* можно было доопределить (продолжить) на все такие множества с сохранением основного свойства счётной аддитивности (2.4.1). Конструкцию, позволяющую удостовериться в возможности такого продолжения, дал французский математик Лебег (мера Лебега). Эта конструкция сложна, и мы не рассматриваем её. Она, однако, обобщается на пространства Ω весьма общего вида, например, на пространства, составленные из функций, что важно для теории случайных процессов.

До сих пор мы говорили об усложнениях, которых безусловно требует аксиоматика Колмогорова, но с ней связаны и важнейшие упрощения. Введение меры, обладающей свойством счётной аддитивности, позволяет воспользоваться понятием интеграла Лебега, который, как концепция, несравненно проще и общее интеграла Римана. Все основные понятия теории случайных величин в общем случае оказываются не длиннее и не сложнее, чем то, что мы уже изложили для дискретного случая. Таким образом, в основных понятиях теории вероятностей возникает замечательная простота, общность и порядок.

Однако, интеграл Лебега не более, чем концепция. Никто не вычисляет интегралов, используя лебеговское продолжение меры, вычисления ведутся с помощью интегралов Римана.

Необходимо сказать, что имеет место некоторая беда, связанная с преподаванием математического анализа в вузах, как в нашей стране, так и за рубежом. В общем, ничего плохого нельзя сказать о той части курса анализа, которая относится к функциям одной переменной, кроме того, что она несколько нудна. Ужас начинается при переходе к функциям нескольких переменных. Трактовка дифференциального, и особенно интегрального исчисления для функций нескольких переменных сейчас совершенно неудовлетворительна. Например, чего стоит набор формул Грина, Стокса, Остроградского, которые преподаются вне всякой связи друг с другом. Между тем, существует единый взгляд на все эти формулы, включающий также и формулу Ньютона – Лейбница. О нём, однако, нельзя прочесть в учебниках математического анализа, зато можно в лекциях Арнольда [6] по теоретической механике.

Интересы изложения теории вероятностей страдают также от неудовлетворительного изложения математического анализа, хотя и не так сильно, как интересы теоретической механики. Например, мы здесь не можем воспользоваться ни понятием интеграла Лебега, ни рядом полезных свойств обычных многомерных интегралов. Ограничимся необходимым минимумом.

Точно так же, как в дискретном случае, от самих случайных величин переходят к их распределениям. Рассуждения, однако, должны быть многомерными, т. е. пригодными не для одной случайной величиной, а сразу для нескольких. Иначе говоря, будем рассматривать вектор $\xi = \xi(\xi_1, \xi_2, ..., \xi_n)$, составленный из случайных величин $\xi_1, \xi_2, ..., \xi_n$. Основной принцип многомерного рассмотрения состоит в том, что нужно вводить такие характеристики, которые допускают лёгкий переход от одной координатной системы к другой. Однако, во многих учебниках вводилась так называемая совместная функция распределения

$$F_{\xi_1,\xi_2,...,\xi_n}(x_1, x_2, ..., x_n) = P(\xi_1 < x_1, \xi_2 < x_2, ..., \xi_n < x_n).$$

При этом переход от одних координат $x_1, x_2, ..., x_n$ к другим $y_1, y_2, ..., y_n$ не только труден, но его, в сущности, даже невозможно записать формулой фактически не вводя вероятностной меры

$$\mu_\xi(A) = \mu_{\xi_1,\xi_2,...,\xi_n}(A) = P\{\xi = (\xi_1, \xi_2, ..., \xi_n) \in A\}.$$

Под знаком вероятности в правой части стоит событие: вектор $\xi = \xi(\xi_1, \xi_2, ..., \xi_n)$ оказался элементом множества A. Поэтому многомерная функция распределения практически бесполезна. На самом деле приходится иметь дело с плотностью распределения

$$p_\xi(x) = p_{\xi_1,\xi_2,...,\xi_n}(x_1, ..., x_n).$$

Она определяется следующим требованием: для любого (не слишком сложного) множества A в многомерном пространстве

$$P\{\xi \in A\} = \int_A \cdots \int p_{\xi_1,\xi_2,...,\xi_n}(x_1, x_2, ..., x_n)\,dx_1 dx_2 ... dx_n.$$

Плотность распределения играет ту же роль, что и распределение случайной величины в дискретном случае. В частности,

$$\mathrm{E}f(\xi_1,\dots,\xi_n)=\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} f(x_1,\dots,x_n)\,p_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n)dx_1\dots dx_n.$$

Важнейшую роль играет формула, связывающая плотности случайного вектора в разных системах координат. Эта формула есть частный случай формулы замены переменных в многократных интегралах, и мы не приводим её; в обычном курсе математического анализа нет даже нужных обозначений. Но из неё немедленно выводятся плотности распределения для суммы, произведения, частного и других операций над случайными величинами.

В случае одномерных величин понятие функции распределения, напротив, весьма полезно. Вот его определение:

$$F_\xi(x) = P(\xi < x),$$

где x любое вещественное число.

Если существует плотность распределения $p_\xi(x)$, то

$$F_\xi(x)=\int_{-\infty}^{x} p_\xi(x)dx.$$

Приведём ещё формулы для ожидания в одномерном случае:

$$\mathrm{E}\xi=\int_{-\infty}^{\infty} xp_\xi(x)dx,\ \ \mathrm{E}f(\xi)=\int_{-\infty}^{\infty} f(x)\,p_\xi(x)dx,$$

а для дисперсии

$$D\xi=\mathrm{E}(\xi-\mathrm{E}\xi)^2=\int_{-\infty}^{\infty}(x-\mathrm{E}\xi)^2\,p_\xi(x)dx.$$

Глава 3. Испытания Бернулли. Присяжные Пуассона

3.1. Испытания Бернулли. Итак, задавать вероятности независимых событий несравненно проще, чем вероятности зависимых событий. Поэтому вероятностные модели с независимыми событиями имеют гораздо больше шансов на практические применения. Простейшей, а потому наиболее широко применимой, является модель, в которой представляют себе некоторое число n независимых опытов, каждый из которых может кончиться одним из двух исходов. Один из них называется *успехом*, другой – *неудачей*. Вероятность успеха считается одинаковой во всех опытах и обозначается через p, а в таком случае вероятность неудачи в каждом опыте есть $q = 1 – p$. Обозначим для краткости успех единицей, неудачу нулём и тогда результат n опытов изобразится последовательностью ω нулей и единиц длины n. Итак, множество элементарных событий Ω = {ω} состоит из всех последовательностей нулей и единиц длины n, т е. всего имеет 2^n элементов. Учитывая независимость результатов отдельных опытов, мы должны дать определение, в силу которого вероятность $p(\omega)$ каждого элементарного события ω подсчитывается по следующему правилу: каждой единице из последовательности надо поставить в соответствие число p, каждому нулю число q и полученные числа перемножить. В результате имеем

$$P(\omega) = p^{\mu(\omega)}q^{n-\mu(\omega)}.$$

Здесь $\mu(\omega)$ есть число единиц в последовательности ω.

Эксперименты, описываемые такой вероятностной моделью, называются испытаниями Бернулли. Случайная величина $\mu = \mu(\omega)$ называется числом успехов в n таких испытаний.

Найдём распределение этой величины. Очевидно, её возможными значениями являются числа 0, 1, ..., n, так что

$$P\{\mu = m\} = \sum P(\omega) = \sum p^{\mu(\omega)}q^{n-\mu(\omega)} =$$
$$\sum p^m q^{n-m} = p^m q^{n-m} \cdot (\text{число таких } \omega, \text{ что } \mu(\omega) = m).$$

Суммирования происходят по ω: $\mu(\omega)= m$. Однако, число последовательностей ω, таких, что $\mu(\omega) = m$, равно, очевидно, числу способов выбрать m символов среди n, т. е. равно C_n^m. Итак, имеем

$$P\{\mu = m\} = C_n^m p^m q^{n-m}. \qquad (3.1.1)$$

Мы получили основную формулу для испытаний Бернулли. Теория, как мы видим, вновь оказывается почти тривиальной, но нетривиально научиться применять её, т. е. научиться находить на практике те явления, которые достаточно хорошо описываются схемой Бернулли.

Классическим примером испытаний Бернулли является бросание монеты. При попытке найти нечто поинтереснее, мы вступаем в область сомнительного. Например, можно ли смотреть на рождение ребёнка того или иного пола как на испытания Бернулли, считая успехом, скажем, рождение мальчика? С точки зрения генетических представлений это достаточно естественно. Но из этих представлений с той же естественностью следует, что рассматриваемая вероятность $p = 1/2$. А на самом деле частота рождения мальчиков несколько больше 1/2, и это установлено на таком большом материале, что невозможно принять гипотезу $p = 1/2$.

Но может быть можно в таком случае принять модель испытаний Бернулли с $p \neq 1/2$? Оказывается, что нет: в испытаниях Бернулли вероятность успеха должна быть постоянной, тогда как статистические данные достоверно показывают, что после крупных войн частота рождения мальчиков увеличивается[3.1]. Обсуждается также вопрос о зависимости вероятности рождения мальчиков от социальных условий, в которых живёт семья. Таким образом, модель испытаний Бернулли в данном случае не полностью соответствует действительности.

Итак, произведя статистическое исследование частоты рождения мальчиков, мы узнаём, что модель испытаний Бернулли, строго говоря, неприменима. Но поскольку частота рождения мальчиков всё-таки очень близка к 1/2, отвергнуть гипотезу о применимости схемы Бернулли можно только путём аккуратного статистического исследования, основанного на глубоких следствиях формулы (3.1.1). В следующей главе мы увидим, как это можно сделать. Применение методов теории вероятностей приводит к выводу, что, строго говоря, мы не имеем права рассуждать о вероятности рождения мальчика (или статистической устойчивости). Но в конечно счёте мы узнаём гораздо больше, чем в том случае, если бы обнаружилось идеальное согласие с теорией вероятностей: мы достоверно узнаём, что существует ещё непознанный механизм, регулирующий численность мужчин и женщин.

Модель испытаний Бернулли часто применяется для оценки тех или иных планов приёмочного контроля. При этом выпуск брака (неудача) или годной продукции (успех) должны описываться схемой Бернулли. После того, что сказано в первой главе о возможности вероятностного описания выпуска брака, становится совершенно ясно, что модель испытаний Бернулли применима лишь в случае хорошо налаженного производства.

Мы подробно разберём попытку применить модель испытаний Бернулли к вероятностям судебных приговоров. Эта попытка связана с именами таких первоклассных учёных, как Лаплас и Пуассон. Её разбор чрезвычайно поучителен, так как показывает на историческом примере, что превосходное владение математическими методами теории вероятностей может сочетаться с абсолютно неверным подходом к действительности.

3.2. Присяжные Пуассона. Лаплас, а вслед за ним Пуассон исследовали вопрос о вероятности ошибки судебного приговора, вынесенного судом присяжных. Естественно, что отдельные присяжные могут ошибаться в своих суждениях. Лаплас наделил присяжных весьма скромной способностью суждений: по его мнению, вероятность ошибки каждого отдельного присяжного есть случайная величина, равномерно распределённая на отрезке [0, 1/2][3.2]. Пуассон не был согласен с этим и полагал, что вероятность правильного суждения должна оцениваться по статистическим данным. Величайшим затруднением для прямой статистической оценки является то простое обстоятельство, что нет возможности установить, правильно или неправильно был осуждён данный обвиняемый[3.3].

Идея Пуассона, которая широко применяется и в настоящее время, заключается в том, что в такой ситуации нужно построить статистическую модель, в которую неизвестная вероятность войдёт в качестве параметра, и попытаться по данным наблюдения определить его. Для этого нужно рассмотреть судебный процесс подробнее. Он основан на материалах следствия. Обозначим через A событие, состоящее в том, что следствием собрано достаточно доказательств, чтобы суд должен был признать обвиняемого виновным, и противоположное событие через $\bar{A}$. Если событие A наступило, то всем присяжным, если бы их суждения были безошибочными, следовало бы единогласно голосовать за обвинение, в противном случае (событие $\bar{A}$ наступило) – за оправдание.

На самом деле из-за ошибок присяжных голоса довольно часто разделяются. Основное предположение Пуассона состоит в том, что разделение голосов происходит согласно схеме испытаний Бернулли. Если n это число присяжных, p, вероятность правильного суждения для каждого отдельного присяжного, то число голосов μ, поданных за осуждение, оценивается следующим образом.

1. Если событие A наступило, μ будет число успехов при n испытаниях Бернулли с вероятностью успеха p. **2.** Если же наступило событие $\bar{A}$, то μ будет числом неудач в тех же испытаниях с вероятностью неудачи $q = 1 – p$.

По французскому законодательству $n = 12$, а обвинительный приговор выносился при $\mu \geq 7$[3.4]. Таким образом, вероятность вынесения обвинительного приговора была равна

$$P_g = P(A)P\{\mu \geq 7/A\} + P(\bar{A})P\{\mu \geq 7/\bar{A}\} =$$

$$P(A)\sum_{m=7}^{12} C_{12}^m p^m (1-p)^{12-m} + [1-P(A)]\sum_{m=7}^{12} C_{12}^m p^{12-m}(1-p)^m. \quad (3.2.1)$$

Судебная статистика даёт частоту вынесения обвинительного приговора, которая приближённо равна P_g, и Пуассон аккуратно проверял устойчивость этой частоты в разные годы. Однако, выражение (3.2.1) содержит два неизвестных параметра: $P(A)$ и p. Зная только P_g, определить их невозможно. Поэтому следует обратиться к статистике, в которой указывалась бы не только, оправдан ли был обвиняемый, но и каким числом голосов. Например, для осуждения ровно семью голосами имеем

$$P_g\{\mu = 7\} = P(A)P\{\mu = 7/A\} + P(\bar{A})P\{\mu = 7/\bar{A}\} =$$
$$P(A)C_{12}^7 p^7 (1-p)^5 + [1-P(A)]C_{12}^7 p^5 (1-p)^7. \quad (3.2.2)$$

Зная левые части соотношений (3.2.1) и (3.2.2), можно в принципе найти неизвестные $P(A)$ и p. Эти левые части приближённо равны частотам, задаваемым судебной статистикой.

Уравнения (3.2.1) и (3.2.2) – высокой степени, и решить их непросто, но Пуассон развивает общий метод для решения таких уравнений, и в конце концов благополучно их решает.

С лёгкой руки Лапласа и Пуассона задачи о вероятностях судебных приговоров вошли во все учебники теории вероятностей XIX века, но в XX веке такие приложения были объявлены абсолютным вздором[3.5]. Нам предстоит разобраться, почему так произошло.

Основным предположением Пуассона была независимость суждений отдельных присяжных. Прекрасно понимая необходимость проверки устойчивости частот, Пуассон [7] ни слова не говорит по поводу экспериментальной проверки независимости. Как можно было бы её проверить? Решая уравнения (3.2.1) и (3.2.2), Пуассон находит, что вероятность правильного суждения отдельного присяжного примерно равнялась 2/3, так что единогласное осуждение имеет вероятность, не превышающую $(2/3)^{12}$, что ниже 0,01 и почти невозможно[3.6]. Между тем, в соседней Англии, как пишет в том же труде Пуассон, закон требует единогласного решения всех 12 присяжных. При этом английские суды выносят гораздо больше обвинительных приговоров (в том числе смертных), чем французские. Напомним, что речь шла о начале XIX века.

Пуассон видит здесь некий предлог для национальной гордости, хотя следовало бы увидеть повод усомниться в собственной вероятностной модели. Справедливости ради следует отметить, что Пуассон всё равно не смог бы проверить свою модель на материалах французской судебной статистики. Охраняя тайну голосования присяжных, французский процессуальный кодекс не требовал от них сообщений, каким числом голосов был вынесен приговор, исключая лишь случай минимального большинства, необходимого для обвинительного приговора.

С современной точки зрения ошибка Пуассона, формально говоря, заключалась в том, что он рекомендовал без дальнейшей проверки вероятностную модель, в которой два неизвестных параметра определялись также по двум величинам, взятым из наблюдений. Ничего не оставалось для проверки статистической модели.

Интересно привести мнение Курно [2] по этому вопросу. Современник Пуассона, он, очевидно, не обладал такой математической мощью, как Пуассон, и тем более, как Лаплас. Но приходится признать, что он в большей степени обладал здравым смыслом, чем эти первоклассные учёные[3.7]. Он отчётливо понимал, что независимость присяжных есть не более, чем предположение, которое должно быть проверено на опыте. Он даже предлагал изменить процессуальный кодекс так, чтобы, не нарушая тайны голосования присяжных, позволило бы получать необходимые статистические данные. По поводу самОй независимости Курно полагал, что если её нет во всей совокупности судебных дел, то во всяком случае можно разделить уголовные дела на группы, в пределах которых будет наблюдаться независимость. Он даже проверял, что если разделить уголовные дела на две группы, преступления против личности и против собственности, то значения параметров $P(A)$ и p, найденные по методу Пуассона, окажутся очень близкими.

В настоящее время мы уверены, что никакой независимости в суждениях отдельных присяжных нет, так что группы, выделяемые Курно, должны были бы скорее всего состоять из одного дела. Но, строго говоря, доказательства такого утверждения не существует. С позиций современной науки точка зрения Курно неуязвима, что лишний раз подтверждает тот факт, что Курно значительно опередил своё время.

Важный для настоящего времени вывод из рассмотренного примера состоит в том, что ни в коем случае нельзя всю наличную статистическую информацию употреблять для определения параметров статистической модели. Совершенно необходимо какую-то часть информации оставлять на проверку самОй модели. В противном случае большие научные усилия могут оказаться потраченными на совершенный вздор.

Глава 4. Содержательные теоремы теории вероятностей

4.1. Теорема Пуассона. В результате работы над исследованием [7], основные выводы из которого мы только что оценили как совершенный вздор, Пуассон тем не менее открыл один из основных теоретико-вероятностных законов. Вычисляя вероятность $P(\mu = m)$ того, что при n испытаниях Бернулли произойдёт m успехов, он нашёл, что при больших n и малых p имеет место приближённая формула

$$P\{\mu = m\} \approx \frac{\lambda^m}{m!} e^{-\lambda}, \qquad (4.1.1)$$

$\lambda = np$; подробнее см. [8]. Точное выражение для $P\{\mu = m\}$ зависит от трёх параметров, n, m и p; в приближённом выражении параметры n и p сведены в один параметр λ.

На первый взгляд это упрощение кажется незначительным, и сам Пуассон не придавал формуле (4.1.1) большого значения. В его работе [7] имелось большое число формул, более точных и почти столь же удобных, как указанная. Правда, сведение двух параметров, n и p, к одному даёт возможность составить для вероятностей (4.1.1) таблицу сравнительно небольшого объёма с двумя входами, m и λ. Для точных выражений $P\{\mu = m\}$ пришлось бы составить таблицу с тремя входами, чего до сих пор не удалось сделать в достаточно удобном виде.

Впрочем, основная роль формулы (4.1.1) заключается не в вычислительном удобстве. Строго говоря, мы эту формулу выражаем в виде математической теоремы [8], относящейся к испытаниям Бернулли, т. е. к независимым испытаниям с двумя исходами с постоянной вероятностью успеха. Но важнейшее обстоятельство состоит в том, что условия этой теоремы можно нарушать, не нарушая её заключения, т. е. равенства (4.1.1).

Например, можно допустить, что разные испытания имеют разные вероятности успеха

$$p_1, p_2, ..., p_n, ...$$

(лишь бы все эти числа были малы). В таком случае точное выражение для $P\{\mu = m\}$ из главы 3, равно как и хорошие приближённые выражения, сразу теряют смысл, потому что они слишком точны. Сравнительно грубое выражение (4.1.1) остаётся справедливым, только вместо λ нужно подставить $(p_1 + p_2 + ... + p_n)$ или среднее арифметическое из этих величин, т. е. среднее значение вероятности успеха.

Заметим теперь, что мы часто лишены возможности повторить эксперимент с данной вероятностью успеха p_i достаточное число раз, так что p_i, может быть, вообще нельзя узнать. Это не обязательно мешает нам узнать среднее из суммы вероятностей.

Итак, приближённое выражение (4.1.1) имеет гораздо больше шансов найти практическое применение, чем точная формула для $P\{\mu = m\}$. Далее, можно ослабить требование независимости отдельных испытаний, можно допустить, что они имеют не обязательно по два исхода, а больше, но такие вопросы выходят за пределы данной книги.

Благодаря такой возможности ослаблять условия теоремы Пуассона без изменения её вывода, распределение Пуассона, придающее значениям $m = 0, 1, 2, \ldots$ вероятности $\frac{\lambda^m}{m!}e^{-\lambda}$, является одним из самых универсальных. Рассмотрим, например, задачу о числе отказов некоторой сложной системы за время T. Пусть система состоит из n элементов, вероятность отказа i-го из которых за время T равно $p_i = p_i(T)$ и в случае отказа повреждённый элемент мгновенно заменяется[4.1]. Тогда число отказов есть число успехов при n испытаниях, причём i-е испытание связано с i-м элементом, а успех этого испытания означает возникновение отказа этого элемента. Но данный элемент может испытать более одного отказа и отказ одного элемента может слегка повлиять на отказы других элементов, однако всё равно, если p_i достаточно малы, то для числа отказов мы ожидаем распределение Пуассона с параметром $\lambda = p_1 + p_2 + \ldots + p_n$. Отклонения возможны, лишь если связи между отказами разных элементов сильны. При малых p_i естественно ожидать следующей зависимости $p_i = p_i(T)$ от T:

$$p_i(T) = p_i' T.$$

Тогда

$$\lambda = \lambda(T) = \lambda' T \qquad (4.1.2)$$

и вероятность m отказов будет равна

$$P\{\mu = m\} \approx \frac{(\lambda' T)^m}{m!} \mathrm{e}^{-\lambda' T}.$$

В частности, для вероятности безотказной работы в течение времени T, т. е. вероятности того, что $\mu = 0$,

$$P\{\mu = 0\} \approx \mathrm{e}^{-\lambda' T}.$$

Это – известный показательный закон для времени безотказной работы. Закон Пуассона и показательный закон согласуются, следовательно, друг с другом. Возникает некоторое гармоничное соответствие, которое можно надеяться с пользой применять к практическим задачам.

В нашей модели мы не учли влияния старения. Но его учёт означает, что линейную зависимость (4.1.2) применять нельзя, её нужно заменить более сложной зависимостью вида

$$\lambda = \lambda(T).$$

Фактически $\lambda(T)$ аппроксимируется функций возможно более простого вида, например, многочленом, коэффициенты которого определяются тем или иным методом, например, методом наименьших квадратов. Но число параметров, которые приходится включать в модель при учёте старения, становится больше, чем без его учёта, и сама модель поэтому будет заслуживать меньшего доверия.

В общем-то, для числа отказов, с учётом старения или без него, естественно применять пуассоновский закон, лишь его параметр определяется по-разному. Конечно, согласие пуассоновского закона с фактическими данными следует проверять статистическими критериями. Если хорошего согласия нет, то, пожалуй, естественнее подозревать статистическую неоднородность имеющихся данных, а не применимость закона Пуассона. Лишь после проверки статистической однородности можно думать о подборе какого-то иного распределения для числа отказов.

Сказанное, конечно, не относится к тому случаю, когда отказы различных элементов сильно связаны. Например, если отказ одной части системы автоматически влечёт отказ другой части, то общее число отказов удваивается. Если даже величина μ имеет распределение Пуассона, то величина 2μ его уже не имеет. В таких случаях для общего числа отказов естественнее говорить о нормальном распределении. И, конечно, распределение Пуассона применимо лишь в том случае, когда отказ в самом деле является редким событием.

Блестящий набор других примеров на применение закона Пуассона может найти в книге Феллера [5]. К этому набору можно добавить разве лишь теорию редких выбросов случайных процессов. Как и все редкие события, число этих выбросов, превышающих достаточно высокий уровень, имеет распределение Пуассона.

4.2. Центральная предельная теорема (ЦПТ)[4.2]. Закон Пуассона определяется единственным параметром λ. Нетрудно показать, что λ является ожиданием случайной величины, распределённой по закону Пуассона. Теперь мы рассмотрим ещё более универсальный закон теории вероятностей, так называемое нормальное распределение. Этот закон определяется двумя параметрами, ожиданием и дисперсией. Это распределение было открыто приблизительно в одно и то же время Гауссом и Лапласом, которые исходили из совершенно различных соображений[4.3].

Гаусс обнаружил, что именно в случае нормального распределения ошибок наблюдения за оценку истинного значения измеряемой величины наиболее естественно брать среднее арифметическое отдельных измерений. Лаплас исходил из открытого им чрезвычайно сильного метода для вычисления вероятностей суммы случайных величин. Идеи Гаусса важны для обработки результатов измерений, и они получили дальнейшее развитие в математической статистике в виде так называемого *метода наибольшего правдоподобия*. Идеи Лапласа относятся к свойствам арифметических операций над большим числом случайных величин и составляют в сущности основу современной теории вероятностей. В данной книге, посвящённой теории вероятностей, а не математической статистике, мы примем подход Лапласа.

Рассмотрим произвольные независимые случайные величины

$$\xi_1, \xi_2, ..., \xi_n, \tag{4.2.1}$$

принимающие одни и те же значения

$$0, \pm 1, \pm 2, ..., \pm m$$

с одними и теми же вероятностями

$$P\{\xi_i = m\} = p_m.$$

Такие величины называются одинаково распределёнными. Эти вероятности p_m могут быть произвольными, подчинёнными лишь условию равенства единице их суммы и условиям достаточно быстрого убывания при $m \to \pm\infty$. Точнее говоря, нужно, чтобы величины ξ_i имели конечное ожидание и конечную дисперсию

$$\mathrm{E}\xi_i = \sum_{m=-\infty}^{\infty} mp_m = a, \ \mathrm{D}\xi_i = \sum_{m=-\infty}^{\infty} (m-a)^2 p_m = \sigma^2.$$

В остальном набор вероятностей$\{p_m\}$ совершенно произволен, и, в частности, может не описываться никаким конечным числом параметров.

Лаплас открыл, что при большом числе слагаемых (4.2.1) распределение их суммы делается несравненно проще распределения отдельных слагаемых. При некоторых дополнительных условиях [8, гл. 8]

$$\sigma\sqrt{n}P\{\xi_1 + ... + \xi_n = m\} \approx \frac{1}{\sqrt{2\pi}}\exp[-\frac{-x_n^2(n)}{2}],$$

$$x_n(m) = \frac{m-na}{\sigma\sqrt{n}}. \tag{4.2.2a,b}$$

Полезно иметь в виду следующие простые соображения, которые помогут понять геометрический смысл равенств (4.2.2). Предположим, что мы собрались графически изобразить распределение каждой из случайных величин (4.2.1) и их суммы.

Для изображения распределения каждого отдельного слагаемого мы нарисуем ось абсцисс, нанесём на ней точки

$$0, \pm 1, \pm 2, ..., \pm m ...$$

и изобразим каждую вероятность p_m прямоугольником с единичным основанием, середина которого лежит в точке m, а площадь (т. е. высота) равна p_m. Получится некоторый, вообще говоря, бессистемный набор прямоугольников. Если теперь на том же чертеже мы попытаемся изобразить распределение вероятностей суммы этих случайных величин при большом n, то нам это не удастся, так как возможные значения такой суммы могут быть очень велики, а вероятности отдельных значений, как можно показать, малы. Потребуется, следовательно, изменить масштаб, т. е. вместо значений $m = (\xi_1 + ...+ \xi_n)$ откладывать некоторые числа вида

$$(1/B_n)(m - A_n).$$

Суть открытия Лапласа теперь можно выразить одной фразой: сдвиг шкалы A_n надо выбрать равным

$$A_n = \mathrm{E}(\xi_1 + ... + \xi_n) = na,$$

а масштабный множитель равным

$$B_n = \sqrt{\mathrm{D}(\xi_1 + ... + \xi_n)} = \sigma\sqrt{n}.$$

Случайная величина

$$s_n^* = \frac{1}{\sqrt{\mathrm{D}(\xi_1 + ... + \xi_n)}}[(\xi_1 + ... + \xi_n) - \mathrm{E}(\xi_1 + ... + \xi_n)] \qquad (4.2.3)$$

называется нормированной суммой. Очевидно,

$$\mathrm{E}s_n^* = 0,\ \mathrm{D}s_n^* = 1.$$

Числа (4.2.2b) являются возможными значениями указанной нормированной суммы. Попытаемся изобразить вероятности $P\{s_n^* = x_n(m)\}$ прямоугольниками, основания которых равны

$$x_n(m+1) - x_n(m) = \frac{1}{\sigma\sqrt{n}},$$

середины оснований находятся в точках $x_n(m)$, а площади равны вероятностям

$$P\{s_n^* = x_n(m)\} = P\{\xi_1 + ... + \xi_n = m\}.$$

Высоты этих прямоугольников должны равняться

$$\sigma\sqrt{n}P\{s_n^* = x_n(m)\} = \sigma\sqrt{n}P\{\xi_1 + ... + \xi_n = m\}.$$

Таким образом, ввиду (4.2.2а) верхние основания этих прямоугольников почти точно лягут на не зависящую ни от чего и раз навсегда вычисленную кривую, уравнение которой

$$y = y(x) = \frac{1}{\sqrt{2\pi}}\exp[-\frac{x^2}{2}]. \qquad (4.2.4)$$

Результат совершенно непредвиденный и почти чудесный! Из беспорядка в вероятностях p_m каким-то образом рождается универсальная кривая, задаваемая соотношением (4.2.4). Это попросту происходит за счёт операции суммирования случайных величин и преобразования масштаба чертежа. Таково замечательное открытие Лапласа, без которого теория вероятностей почти не имела бы оригинального, т. е. не сводящегося к известным понятиям математического анализа, содержания.

Правда, утверждение (4.2.4) имеет некоторые исключения. Например, если все случайные величины (4.2.1) принимают только чётные значения (т. е. если $p_m \neq 0$ только для чётных значений m), то сумма (ξ_1 + ...+ ξ_n) также всегда чётна, при нечётных же m левая часть (4.2.2а) равна нулю и это равенство нарушается. Это исключение, по существу единственно [8, гл. 8]. Но для того, чтобы избавиться от него, и по ряду других причин ЦПТ предпочитают формулировать в терминах функций распределения. Формулировка этой теоремы, по существу известная уже Лапласу, такова.

Теорема. Пусть ξ_1, ..., ξ_n, ... будет последовательностью независимых одинаково распределённых случайных величин, имеющих конечные ожидание a и дисперсию σ^2, а s_n^*, см. (4.2.3), является нормированной суммой этих величин. Тогда, при $n \to \infty$,

$$P\{s_n^* < x\} \to \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\exp[-\frac{x^2}{2}]dx$$

равномерно по всем x, $-\infty < x < \infty$.

Эта теорема ещё менее чувствительна к нарушению её условий, чем даже замечательная в этом смысле теорема Пуассона. Развитие теории вероятностей в значительной мере связано с усовершенствованием доказательства и ослаблением условий этой теоремы. Можно *отрицать*, т. е. заменять менее ограничительным, каждый термин из условия этой теоремы, не уничтожая её заключения. Так, требование одинакового распределения случайных величин отрицал Ляпунов. В результате возникла теорема Ляпунова [8, гл. 8]. Требование независимости отрицал Бернштейн (1926). В последнее время сделана попытка в значительной мере отбросить условие случайности величин (4.2.1). Можно также заменить случайные величины, принимающие числовые значения, на случайные элементы некоторых групп и вместо сложения рассматривать соответствующую групповую операцию. Здесь достигнуты определённые успехи, но в этой книге о них ещё рано говорить.

Нельзя отрицать условие конечности дисперсии [9], ибо в таком случае теорема о сходимости к нормальному закону неверна. Возможно лишь некоторое ослабление этого условия.

4.3. Нормальное распределение. Итак, наиболее широко применимым является распределение Гаусса – Лапласа, плотность которого задаётся формулой (4.2.4). Иными словами, говорят, что случайная величина ξ имеет стандартное нормальное распределение, если для практически любого множества *A* справедливо равенство

$$P\{\xi \in A\} = \int_A \frac{1}{\sqrt{2\pi}} \exp[-\frac{x^2}{2}]dx.$$

Наряду со случайной величиной ξ естественно рассматривать случайные величины

$$\eta = \sigma\xi + a,$$

например, если ξ результат некоторого измерения, а η результат того же измерения в другой системе единиц. Нетрудно показать, что плотность распределения η равна

$$p_\eta(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(x-a)^2}{2\sigma^2}].$$

Ожидание и дисперсия η равны $E\eta = a$ и $D\eta = \sigma^2$. Распределение случайной величины η называется нормальным с параметрами *a* и σ и обозначается N(*a*, σ).

Важнейшее значение имеет следующая теорема: если η_1 и η_2 независимые случайные величины, имеющие распределения $N(a_1, \sigma_1)$ и $N(a_2, \sigma_2)$, то их сумма имеет распределение $N(a_1 + a_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

4.4. Теорема Муавра – Лапласа. Рассмотрим *n* испытаний Бернулли с вероятностью *p* успеха в каждом отдельном испытании. Число успехов μ в *n* испытаниях может быть представлено в виде суммы

$$\mu = \mu_1 + \mu_2 + \ldots + \mu_n,$$

в которой случайная величина μ_k равна числу успехов в *k*-м испытании, т. е. равна 1 или 0 в зависимости от того, был ли в этом испытании успех или нет.

Поскольку все испытания независимы, то μ есть сумма независимых случайных величин и к ней применима ЦПТ. Поэтому распределение величины μ приблизительно нормально с параметрами, как нетрудно подсчитать,

$$\mathrm{E}\mu = np,\ \sqrt{\mathrm{D}«} = \sqrt{npq}\ .$$

В этом и состоит теорема Муавра – Лапласа[4.4].

4.5. Применения центральной предельной теоремы.

Проверка статистической однородности. В первой главе мы много говорили о том, что условием научного применения теории вероятностей является проверка статистической однородности. Теперь, наконец, можно объяснить, каковы основные методы, применяемые для этой цели. Обычно речь идёт о следующей задаче. Произведены n_1 опытов, причём событие A наступило μ_1 раз. Кроме того, произведены другие n_2 опытов, в которых то же событие наступило μ_2 раз. Можно ли думать, что вероятности успеха в обеих сериях опытов одинаковы?

Иначе говоря, является ли различие частот μ_1/n_1 и μ_2/n_2 достаточно малым, чтобы его можно было бы объяснить чисто случайными причинами[4.5]? Естественно думать, что величины μ_1 и μ_2 распределены примерно нормально, безразлично, зависимы ли отдельные опыты или независимы. Если p_1 и p_2 вероятности события A в этих сериях, то

$$\mathrm{E}(\mu_1/n_1) = p_1,\ \mathrm{E}(\mu_2/n_2) = p_2,$$

и, если $p_1 = p_2 = p$, то $\mathrm{E}(\mu_1/n_1 - \mu_2/n_2) = 0$. Далее, естественно считать, что опыты первой и второй серии независимы друг от друга. Поэтому величина $(\mu_1/n_1 - \mu_2/n_2)$ должна иметь приблизительно нормальное распределение с нулевым ожиданием и дисперсией

$$\mathrm{D}(\frac{\mu_1}{n_1} - \frac{\mu_2}{n_2}) = \mathrm{D}(\frac{\mu_1}{n_1}) + \mathrm{D}(\frac{\mu_2}{n_2}).$$

Если члены в правой части известны, при помощи таблиц нормального распределения мы могли бы сказать, можно ли объяснить указанную разность чисто случайными причинами или нет. И вот, для вычисления этих дисперсий (но совсем не для применения ЦПТ), нужно допустить, что опыты в этих сериях независимы друг от друга, т. е., что сделаны две серии испытаний Бернулли (ЦПТ не требует полной независимости). В этом предположении, если $p_1 = p_2 = p$ (это p неизвестно)

$$\mathrm{D}(\frac{\mu_1}{n_1}) + \mathrm{D}(\frac{\mu_2}{n_2}) = p(1-p)(\frac{1}{n_1} + \frac{1}{n_2}). \qquad (4.5.1)$$

Можно показать, что для наших целей неизвестное p в формуле (4.5.1) допустимо заменить на

$$\hat{p} = \frac{\mu_1 + \mu_2}{n_1 + n_2}.$$

Итак, если $p_1 = p_2$. то

$$\xi = (\frac{\mu_1}{n_1} - \frac{\mu_2}{n_2}) \div \sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}$$

имеет приблизительно стандартное нормальное распределение. Значение величины ξ может быть вычислено только лишь исходя из чисел (μ_1, n_1) и (μ_2, n_2), т. е. из известных результатов экспериментов. Как известно, значения ξ, превышающие по абсолютной величине два или три, маловероятны. Получив значение такого порядка, мы должны были бы заключить либо, что гипотеза $p_1 = p_2$ неверна, либо тем не менее верна, но произошло маловероятное событие.

Как выбрать между этими заключениями? Некоторые думают, что здесь может помочь теория статистических решений, которая будто бы позволяет выразить числом риск каждого из возможных решений. Но в этой теории риск выражается через величины, которые либо вовсе не имеют смысла, либо, во всяком случае, никогда не могут быть известны исследователю. Теория, основанная на количественном выражении риска, во всех случаях кроме азартных игр бесполезна[4.6].

На самом деле выбор между двумя указанными решениями является сложной процедурой, которую в настоящее время нет возможности исследовать в рамках математической теории. Приходится каким-то образом сопоставлять важность отрицательных последствий каждого возможного решения, если оно окажется ошибочным. При этом участвуют и научные, и моральные соображения, обозначаемые словом *совесть*[4.7].

Примерно таким же образом, но несколько проще проверяется гипотеза о том, что вероятность успеха p в заданной серии испытаний Бернулли равна данному числу p_0. Если эта гипотеза верна также и в нашей серии n испытаний, из которых μ успехов, то

$$(\frac{\mu}{n} - p_0) \div \sqrt{\frac{p_0(1-p_0)}{n}}$$

имеет приблизительно стандартное нормальное распределение. Именно вычислением указанной величины может быть проверена (и отвергнута) гипотеза о том, что вероятность рождения мальчика равна половине, см. § 3.1.

Преобразование арксинуса. Мы только что изложили идейную сторону вопроса о проверке равенства вероятностей в двух сериях испытаний Бернулли. Теперь мы хотим показать на примере, какие удобные приёмы разработаны для этой цели в математической статистике. В идейном отношении здесь нет ничего нового, но практические удобства велики. Выберем для примера так называемое *преобразование арксинуса*, открытое знаменитым английским статистиком Фишером. Его идея очень проста. Рассмотрим вместо частоты успеха μ/n некоторую функцию $f(\mu/n)$. Имеем

$$f(\frac{\mu}{n}) = f[(\frac{\mu}{n} - p) + p] = f(p) + f'(p)(\frac{\mu}{n} - p) + \dots$$

При большом n эта частота близка вероятности p, поэтому мы пренебрежём остальными членами этого разложения. В этом приближении

$$\mathrm{D}f(\frac{\mu}{n}) = \mathrm{D}[f'(p)(\frac{\mu}{n} - p)] = [f'(p)]^2 \mathrm{D}\frac{\mu}{n} = [f'(p)]^2 \frac{p(1-p)}{n}.$$

Выберем функцию f таким образом, чтобы выражение для дисперсии $f(\mu/n)$ не зависело бы от неизвестного параметра p. Именно, положим, что

$$[f'(p)]^2 p(1-p) = 1.$$

В качестве $f(p)$ можно взять любое решение этого дифференциального уравнения, в частности

$$f(p) = 2\arcsin\sqrt{p}.$$

Поскольку в рассматриваемом приближении $f(\mu/n)$ есть линейная функция от μ, а при больших n распределение μ приблизительно нормально, то выражение

$$f(\frac{\mu}{n}) = 2\arcsin\sqrt{\frac{\mu}{n}} \qquad (4.5.2)$$

распределено приблизительно нормально. Его ожидание примерно равно

$$f(p) = 2\arcsin\sqrt{p},$$

а дисперсия приблизительно равна $1/n$, поскольку именно так мы выбирали функцию f. Таким образом, дисперсия не зависит от неизвестного параметра p. Кроме того, оказывается, что распределение (4.5.2) ещё ближе к нормальному, чем распределение самогО числа успехов μ.

Пусть теперь даны две такие серии испытаний Бернулли, что в них числа испытаний и успехов и вероятности успехов равны n_1,

μ_1, p_1 и n_2, μ_2, p_2, и мы желаем проверить гипотезу равенства этих вероятностей. Образуем величину

$$2\arcsin\sqrt{\mu_1/n_1} - 2\arcsin\sqrt{\mu_2/n_2}\ .$$

При независимости серий она распределена приблизительно нормально, а её ожидание равно

$$2\arcsin\sqrt{p_1} - 2\arcsin\sqrt{p_2}$$

и равно нулю, если верна сформулированная гипотеза. Дисперсия же этой случайной величины есть сумма дисперсий $(1/n_1) + (1/n_2)$, и

$$[2\arcsin\sqrt{\mu_1/n_1} - 2\arcsin\sqrt{\mu_2/n_2}\,] \div \sqrt{1/n_1 + 1/n_2} \qquad (4.5.3)$$

имеет стандартное нормальное распределение N(0, 1). Вычислив значение указанной величины, мы можем принять или отвергнуть гипотезу о равенстве вероятностей.

Математическая статистика полна такими простыми, но весьма удобными приёмами. В данном случае настоящее удобство получается, конечно, при использовании таблицы функции $2\arcsin\sqrt{x}$, которая имеется во многих сборниках статистических таблиц, хотя, впрочем, можно пользоваться и логарифмической линейкой.

Информация о поведении суммы независимых случайных величин. Если мы рассматриваем независимые одинаково распределённые случайные величины (4.5.1) с ожиданием и дисперсией a и σ^2, то их нормированная сумма

$$s_n^* = \frac{\xi_1 + \ldots + \xi_n - na}{\sigma\sqrt{n}}$$

не может при каждом фиксированном n принимать особенно большие значения. Например, принимая нормальное приближение для распределения нормированной суммы, легко находим по таблицам нормального закона, что

$$P\{|\, s_n^* \,| \ \le\ 3\} \ =\ 0{,}997\ ,$$

и почти достоверно, что

$$|\xi_1 + \ldots + \xi_n - na| \le 3\sigma\sqrt{n}. \qquad (4.5.4)$$

Внимание читателя следует, правда, обратить на то, что это утверждение получено нами для любого фиксированного значения n. Если рассматривать последовательность $s_1^*, \ldots, s_n^*, \ldots$, то, конечно, нельзя утверждать, что все случайные величины, входящие в неё, меньше трёх по абсолютной величине. Распределение максимального значения $|s_k|$, $1 \le k \le n$, есть особый вопрос, который мы не будем рассматривать. Но при каждом фиксированном n отклонение суммы $\xi_1 + ... + \xi_n$ от её ожидания na возможно лишь на величину порядка $\sqrt{n}$. Итак, неслучайная величина na играет преобладающую роль по сравнению со случайными колебаниями порядка $\sqrt{n}$, если, конечно, $a \ne 0$.

При делении (4.5.4) на n получаем

$$\left|\frac{\xi_1 + ... + \xi_n}{n} - a\right| \le \frac{3\sigma}{\sqrt{n}}. \qquad (4.5.5)$$

Это неравенство выполнено с вероятностью 0,997, если, конечно, верить в нормальное приближение. Если вместо 3σ взять 4σ или 5σ, то это неравенство будет выполнено с ещё более близкой к единице вероятностью. Итак, при большом n практически достоверно, что разность в левой части (4.5.5) мала по абсолютной величине. Мы получили знаменитый закон больших чисел (ЗБЧ).

Любопытно остановиться на истории его доказательства и истолкования. Давно замечено, что результаты отдельных наблюдений, будь то физические, метеорологические, демографические или иные, колеблются сильно, тогда как средние из их большого числа обнаруживают замечательную устойчивость. Первые статистики видели в этой устойчивости вмешательство провидения, но по мере установления научного взгляда на мир в этом стали видеть научный факт.

В XVIII веке, *веке разума*, очень верили в математику, полагая, что из неё могут быть выведены основные законы не только естественных наук, но также экономики, морали и политики. Поэтому установилось желание смотреть на факт устойчивости средних как на математическую теорему. Это мнение держалось и в XIX веке, и именно так Пуассон истолковывал открытую им форму ЗБЧ. Он полагал, что ему удалось доказать, что средние из реальных наблюдений должны быть устойчивы.

Математическую форму ЗБЧ очень сильно усовершенствовал Чебышёв. Он свёл его доказательство к применению неравенства [Бьенеме-]Чебышёва. Это доказательство можно найти в любом учебнике теории вероятностей. После Чебышёва ЗБЧ стал рассматриваться как очень простая теорема, независимая от ЦПТ и излагаемая в курсах теории вероятностей до ЦПТ. Студентов теперь даже учат пользоваться неравенством

$$P\left[\left|\frac{\xi_1 + ... + \xi_n}{n} - a\right| > \varepsilon\right] \le \frac{\sigma^2}{n\varepsilon^2}$$

для оценки вероятности того, что среднее уклонится от a больше, чем на ε. Впрочем, использование этого неравенства [Бьенеме-]Чебышёва для такой оценки совершено нелепо, поскольку применение ЦПТ обеспечивает гораздо более точный результат. Правда, ЗБЧ в форме Чебышёва требует меньших математических ограничений на случайные величины ξ_i, чем ЦПТ. Но на практике всё равно невозможно проверить, выполнены ли соответствующие математические ограничения, и уж конечно, невозможно отличить случай, когда действует лишь теорема Чебышёва от случая, когда действует и она, и ЦПТ.

Эволюция взгляда на естественнонаучное значение ЗБЧ связано с именем Мизеса. Он обращал особое внимание на то, что никакого математического доказательства того, что среднее из результатов эксперимента, т. е. среднее из ξ_i, должно быть близко к некоторому числу, не может быть. В настоящее время мы гораздо меньше верим в то, что законы внешнего мира можно вывести из математики, чем верили во времена Лапласа и Пуассона. Существует слишком много причин, которые могут сделать ход эксперимента иным, чем следовало бы согласно нашей математической модели. Например, в условия всех известных математических теорем о ЗБЧ входит предположение, что (4.5.1) есть последовательность случайных величин. Практически это требование означает, что можно говорить о вероятности события, состоящего в том, что случайная величина ξ_1 приняла значение из некоторого числового множества A_1, ξ_2, из A_2, и т. д., так что для любых множеств A_1, ..., A_n событие $\{\xi_1 \in A_1, ..., \xi_n \in A_n\}$ должно обладать статистической устойчивостью. Однако, возможных множеств A_1, ..., A_n так много, что нет никакой возможности экспериментально проверить статистическую устойчивость всех таких событий. Нарушение же статистической устойчивости вполне обесценивает любую теорему теории вероятностей и может быть причиной наблюдаемого нарушения устойчивости средних в эксперименте.

Таким образом, естественнонаучное значение ЗБЧ сейчас сводят к тому, что соответствующие теоремы отражают экспериментальный факт устойчивости средних в рамках применимости моделей теории вероятностей.

В первой главе мы отмечали, что существует много вопросов, например, в геологии или экономике (число примеров можно без всякого труда умножить), когда не имеет смысла говорить о статистически однородном ансамбле экспериментов. Любопытно, что в этих условиях также довольно часто имеет место устойчивость средних. Приходится признать, что в настоящее время в рамках какой-то математической модели в таких случаях нет сколько-нибудь удовлетворительного отражения устойчивости средних.

В XX веке исследование ЗБЧ в рамках модели пространства элементарных событий было существенно продвинуто. Именно, был открыт так называемый усиленный ЗБЧ, связанный с именами Бореля и особенно Колмогорова. Чтобы объяснить суть дела, вспомним, что в колмогоровской модели случайные величины (4.5.1) суть функции $\xi_1(\omega)$, ..., $\xi_n(\omega)$ на пространстве элементарных событий. Можно рассмотреть событие, т. е. множество

$$\{\omega : \lim \frac{\xi_1(\omega)+\ldots+\xi_n(\omega)}{n} = a\},\ n \to \infty,$$

состоящее из тех элементарных событий ω, для которых этот предел средних арифметических существует и равен *a*. Теоремы типа усиленного ЗБЧ утверждают, что вероятность этого множества равна единице. Обычный ЗБЧ вовсе не имеет дела с этим множеством, а лишь с множествами вида

$$\{\omega : \left|\frac{\xi_1(\omega)+\ldots+\xi_n(\omega)}{n} - a\right| > \varepsilon\},\ \varepsilon > 0$$

при каждом конечном *n* и утверждает, что вероятности таких множеств стремятся к нулю при $n \to \infty$.

Рассмотрение довольно тонких математических примеров показывает, что усиленный ЗБЧ действительно усиленный. Если он имеет место, то наверняка имеет место и обычный закон, если же имеет место обычный закон, то ещё не обязательно выполняется усиленный.

Из предложений об усиленном ЗБЧ отметим весьма изящную теорему Колмогорова, согласно которой для независимых одинаково распределённых случайных величин достаточно существования ожидания, чтобы был выполнен усиленный ЗБЧ. Математический интерес этого утверждения состоит в том, что не требуется существования дисперсий.

Доказательство этой теоремы потребовало создания специфического математического аппарата, и в частности Колмогоров нашёл замечательное *неравенство Колмогорова*, которое обобщало неравенство [Бьенеме–]Чебышёва. Сам аппарат, несомненно, может иметь естественнонаучные применения. Однако, таких применений, когда из одной формулировки усиленного ЗБЧ можно было бы извлечь больше, чем из обычного ЗБЧ, не найдено.

Это связано с тем фактом, что, как отмечалось во второй главе, случайная величина ξ = ξ(ω) как функция от элементарного события ω обычно не наблюдается. Мы знаем значение ξ(ω), но не знаем самогО ω. Скорее можно говорить о вероятностях различных событий. Аналогично, довольно бессмысленно говорить о наблюдении предела среднего $\bar{\xi}$, мы можем говорить лишь о наблюдении этого среднего при конечном *n*. Указанные обстоятельства делают маловероятными какие-либо нематематические применения усиленного ЗБЧ.

В заключение вопроса о применениях ЦПТ остановимся на *применении*, данном ни кем иным, как бесспорно великим и гениальным учёным, Лапласом, но для нас интересным лишь как историко-психологический курьёз. Лаплас открыл уже отмеченный нами факт, что при больших *n* сумма $\xi_1 + \ldots + \xi_n$ ведёт себя примерно как неслучайная величина *na*, в то время, как случайные колебания имеют порядок $\sqrt{n}$, так что с ростом *n* эта неслучайная величина в конце концов возобладает над случайными колебаниями. В частности, если $a > 0$, то и сумма случайных величин будет положительна.

Без особых объяснений Лаплас делает отсюда вывод, что если далеко за морем некое государство имеет колонию, то в конце концов она добьётся независимости. Стремление колонии освободиться Лаплас очевидно представлял себе как некоторый случайный фактор, действие которого накапливается со временем, тогда как усилия метрополии удерживать колонии в подчинении он понимал как случайные величины с нулевым ожиданием. Если первое остаточно понятно, то второе весьма странно[4.8]. Впрочем, в конечном счёте Лаплас был прав: колонии действительно освободились. Но стремление удерживать колонии очевидно нельзя было рассматривать как случайную величину: оно не обладает статистической устойчивостью[4.9]. Если в XIX веке ничего не стоило послать экспедиционный корпус для подавления восстания в колонии, то в наше время это вызвало бы энергичный протест и в метрополии.

Очевидно, что учёный, открывший нечто замечательное, как Лаплас – ЦПТ, не может удержаться, чтобы не применять своё открытие во всех случаях жизни. Например, в наше время Винер предлагал применять теорию экстраполяции случайных процессов для прогноза полёта самолёта в связи с зенитной стрельбой. Но траектория самолёта – не случайный процесс, по крайней мере, не такой случайный процесс, для каких существует теория экстраполяции, и предложение Винера не имеет никакого смысла.

Наука есть, очевидно, коллективное творчество. Правда, не бесспорно, можно ли сделать коллективно крупное открытие, или же обязательно требуется, чтобы в коллективе был выдающийся учёный, на которого, в сущности, работают другие члены коллектива. Зато освобождение от того вздора, который некоторые крупные учёные имеют обыкновение присовокуплять к своим настоящим открытиям, есть бесспорно коллективный процесс.

4.6. Когда нельзя применять центральную предельную теорему. ЦПТ служит одним из оснований для того, чтобы думать, что результаты наблюдений обычно имеют нормальное распределение. Если мы не знаем параметров соответствующего закона, а имеем лишь сами наблюдения $\xi_1, ..., \xi_n$, то существуют способы приближённо определять параметры. В соответствии с ЗБЧ

$$a = \mathrm{E}\xi_i \approx (1/n)\,(\xi_1 + ... + \xi_n) = \bar{\xi}.$$

Можно показать, см., например, [8], что

$$\sigma^2 \approx \frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2 = s^2 .$$

Разработана так называемая теория ошибок, см. например, там же[4.10], которая позволяет определить точность найденных таким образом приближённых значений этих величин.

В общем, наблюдения обычно неплохо описываются найденным таким образом нормальным распределением. Иными словами, если $F(x) = P\{\xi_i < x\}$, а через $N(x, \bar{\xi}, s)$ обозначить соответствующую вероятность, найденную из нормального распределения, то

$$F(x) \approx N(x, \bar{\xi}, s).$$

Но это приближённое равенство иногда нарушается весьма чувствительным образом. Речь идёт о значениях x, отвечающих значениям $F(x)$, близким к 0 или 1. Это так называемые *хвосты* распределения $F(x)$. Рассмотрим сначала, почему *хвосты* имеют особое практическое значение.

Предположим, что мы собираемся построить какое-нибудь сооружение, например, Останкинскую телебашню, которому предстоит выдерживать значительные ветровые нагрузки, или, если угодно, водосбросное сооружение, которое должно пропускать паводковые воды и т. д. Мы стремимся рассчитать своё сооружение на такую скорость ветра, которая встречается достаточно редко, например, раз в сто лет. Но как узнать, какова эта скорость? Иными словами, если $\xi(t)$ скорость ветра в момент t, то требуется указать такое число x, что

$$P\{\max \xi(t) \geq x\} = 0{,}01,\ 0 \leq t \leq 1,$$

где t измеряется в годах и левая часть неравенства является наивысшей годичной скоростью ветра.

Допустим, что известны значения $\xi_1, \xi_2, ..., \xi_n$ максимальной скорости ветра за первый, второй, …, n-й год, в течение которых велись метеорологические наблюдения. Правда, сплошные записи скорости ветра не велись, скорость измерялась несколько раз в день, так что искомые скорости в сущности неизвестны. Но пока отвлечёмся от этого крайне существенного затруднения.

Итак, имеются наблюдения случайной величины ξ, наивысшей скорости ветра за год, и мы хотим назначить такое x, чтобы

$$P\{\xi \geq x\} = 0{,}01. \qquad (4.6.1)$$

Будь число n очень велико, то x следовало бы выбрать таким образом, чтобы примерно сотая часть из величин ξ_i превосходила x. Но беда в том, что число n, т. е. число лет, за которые имеются наблюдения, гораздо меньше ста. И если x таково, что выполнено равенство (4.6.1), т. е. если

$$P\{\xi_i \geq x\} = 0{,}01 \text{ ля каждого } i,$$

то число величин ξ_i, превосходящих x, подчиняется закону Пуассона с параметром $\lambda = 0{,}01n < 1$. Поэтому скорее всего все наши наблюдения ξ_i меньше x, так что мы можем только сказать, что x должно быть больше каждого ξ_i. У нас нет никакой оценки для x сверху.

Возникает искушение *сгладить* наши наблюдения ξ_1, ..., ξ_n каким-либо законом, например, нормальным $\mathrm{N}(x;\overline{\xi},s)$ и определить x из уравнения

$$\mathrm{N}(x;\overline{\xi},s)=1-0,01=0,99.$$

Иначе говоря, предлагается считать, что *хвосты* неизвестной функции $F(x)$ совпадают с *хвостами* нормального закона. Мы хотим обратить внимание читателя на то, что такая процедура не заслуживает доверия как при использовании нормального, так и любого другого закона. Для этого утверждения имеются и теоретические, и опытно-статистические основания.

Теоретические основания заключаются в том, что ЦПТ утверждает лишь, что разность между точной функцией распределения $P\{s_n^* < x\}$ и нормальным законом мала:

$$P\{s_n^* < x\}-\mathrm{N}(x)\to 0.$$

К примеру, если P = 0,95, $\mathrm{N}(x)$ = 0,99 то разность, равная 0,04, достаточно мала. Однако, относительная ошибка

$$[1-P\{s_n^* < x\}]\div[1-\mathrm{N}(x)]=400\%$$

весьма велика. Нам небезразлично, что фактически $P\{s_n^* \ge x\}$ = 0,05, так что событие $\{s_n^* \ge x\}$ происходит в одном случае из 20 (так сказать, раз в двадцать лет), тогда как с помощью нормального приближения мы находим, что оно происходит в одном случае из ста (раз в сто лет). Подчёркиваем, что ЦПТ не утверждает, что

$$[1-P\{s_n^* < x\}]\div[1-\mathrm{N}(x)]\to 1 \qquad (4.6.2)$$

равномерно по x, и такое утверждение на самом деле неверно.

Итак, в области вероятностей, близких к 1 (и к 0), использование нормального приближения может дать (и, как правило, действительно даёт) большую относительную ошибку, в то время как в согласии с ЦПТ абсолютная ошибка будет мала. В частности, нужно иметь в виду, что равенство

$$P\{\left|\frac{\xi_1+...+\xi_n}{n}-a\right| \le \frac{3\sigma}{\sqrt{n}}\} = 0,997,$$

использованное в § 4.5, несколько условно. Вместо 0,997 легко может быть 0,990 или 0,980, и только при очень больших n

действительно будет 0,997. Возможно также, что вместо 0,997 будет просто 1.

Отношение в левой части соотношения (4.6.2) исследуется в теории вероятностей с помощью так называемых *теорем о больших уклонениях* [5]. Но практическое значение таких теорем недостаточно ясно, хотя, впрочем, они показывают, что результат не улучшится, если вместо нормального распределения применять другие, часто встречающиеся законы (например, пирсоновские).

Что касается статистического опыта работы с *хвостами* распределений, то он показывает, что поведение *хвостов* нерегулярно. Возможно, что на них особое влияние оказывают нарушения статистической однородности, которые воздействовали на исход отдельных опытов. В таких случаях проблема описания *хвостов* распределения статистическими методами безнадёжна.

Проблема описания возможных *раз в сто лет* значений скорости ветра осложняется ещё тем, что речь идёт о значениях скорости ветра, максимальных за год. Если значение скорости ветра в данный момент считать нормально распределённым, то максимальное значение за год естественно было бы рассматривать с помощью какого-то распределения для экстремальных значений. Но эти распределения выведены лишь для независимых величин и не могут учесть эффект постепенного нарастания скорости ветра в определённых метеорологических условиях. Кроме того, сама теория экстремальных значений часто прилагается к фактическим данным с большими натяжками.

Если вспомнить ещё об отсутствии непрерывных записей скорости ветра, то можно совершенно твёрдо сказать, что никакого научного способа узнать, насколько сильным может быть ветер *один раз в сто лет*, мы в настоящее время не имеем. Проектировщикам сооружений следовало бы найти какой-либо иной способ для утверждений, какая надёжность заложена в их сооружениях.

Библиография автора

1. А. Н. Колмогоров. *Основные понятия теории вероятностей*. М. – Л., 1936.

2. О. Курно. *Основы теории шансов и вероятностей*. М., 1970. Перевод с французского издания 1843 г.

3. Р. фон Мизес. *Вероятность и статистика*. М. – Л., 1930.

4. Р. Фейнман, Р. Лейтон. М. Сэндс. *Фейнмановские лекции по физике*, вып. 1. М., 1967.

5. В. Феллер. *Введение в теорию вероятностей и её приложения*, т. 1 – 2. М., 1967.

6. В. И. Арнольд. *Лекции по классической механике*, вып. 1 – 2. М., 1968.

7. S.-D. Poisson. *Recherches sur la probabilité des jugements* etc. Paris, 1837, 2003, **S, G,** 52.

8. Б. В. Гнеденко. *Курс теории вероятностей*. М., 1969.

9. Б. В. Гнеденко. А. Н. Колмогоров. *Предельные распределения для сумм независимых случайных величин*. М. – Л., 1949.

Примечания

1. Сложный для своего времени математический аппарат применили и Бейес, и Муавр.

1.1. Можно было сослаться на неустойчивые движения, чувствительные к небольшим изменениям начальных условий. Громадным, и притом качественным обобщением этого явления является хаотическое движение, которое усиленно изучается быть может в последние 70 лет.

1.2. Советские статистики не смели всерьёз изучать статистику населения. Слишком серьёзный урон нанесла населению зверская сталинская политика.

1.3. Лишь в конце брошюры (Прим. 4.9) автор определил устойчивость, да и то лишь косвенно.

1.4. Это, правда, лишь мельком упомянутое понятие, не разъяснено, хотя выше автор определил понятие *подмножества* (которое мы не сочли нужным повторять). Подобное противоречие характерно для научно-популярной литературы.

2.1. Классическое определение вероятности восходит к Муавру (а не к Лапласу). Точнее, это не определение, а формула для вычисления вероятностей, область применения которой весьма ограничена. Мало того, она основана на порочном круге.

2.2. Автор ограничился случаем квартир двух видов. Описание жеребьёвки в Талмуде (Шейнин 1998) указывает, что её участники не доверяли независимости её результатов от очерёдности извлечения билетиков. Лаплас (1812/1886, с. 413) был, видимо, первым, заметившим, что жеребьёвка в два этапа стремится уравнять шансы её участников.

2.3. О трудности выравнивания шансов извлечения билетиков из урны см. Fienberg (1971).

2.4. Субъективные вероятности приходится учитывать при оценке экспертных мнений, и их же применяют присяжные. Якоб Бернулли ввёл неаддитивные субъективные вероятности, идею которых он мог отыскать в средневековом учении пробабилизма (мнение каждого Отца церкви считалось вероятным). Наконец, о машинной диагностике см. *Интеллектуальные … системы* (2016).

2.5. Эта рекомендация почти бесполезна. Осторожность нужна даже при простейших вычислениях.

3.1. Это явление признаётся и называется *феноменом военных лет*. Заметно и некоторое изменение соотношения полов новорождённых от одной страны к другой, которое, впрочем, сильно связано с упомянутым феноменом. См. в Википедии данные за 2014 г.

3.2. Лаплас (Шейнин 2019, § 8.1-1)оценивал ошибки присяжных в соответствии с бейесовским подходом с постоянными априорными вероятностями, но нелепо предполагать, что он не представлял, что его выводы окажутся грубо приближёнными. В другом месте (там же, § 8.2-1) он указал, что следует выявлять истину постепенно, по мере накопления данных.

3.3. Пуассон (1837) даже рекомендовал отказаться от понятий *виновен/не виновен* и принять взамен *подлежит наказанию*/*оправданию* в зависимости от собранных данных.

3.4. Основной целью Пуассона (1837) было определение оптимального соотношения голосов присяжных при обвинении (7 против 5 или 8 против 4), и он мог надеяться, что влияние взаимной зависимости присяжных окажется при этом неизменным. Его критерием был минимум неизбежных судебных ошибок обоих видов (осуждение невинного и оправдание виновного). Gelfand & Solomon (1973, с. 273) несколько смягчили проблему взаимной зависимости присяжных.

3.5. В XIX веке рассуждения Пуассона не принимались всерьёз, но в XX веке проявился интерес к ним. Вообще же уже Лейбниц (Шейнин 2019, § 4.1.2) считал, что чрезвычайно важно учитывать все обстоятельства уголовного дела. Пуанкаре (1896/1999, с. 410) лишь заявил, что присяжные ведут себя как панургово стадо (т. е. слепо следуют за кем-то из них), а Бернштейн (1926/1945, с. 410), несомненно, как и Пуанкаре, не зная существа дела, посчитал, что Лаплас и Пуассон явно ошибались.

3.6. Вывод автора не вполне убедителен. Взяв 4/5 вместо 2/3, мы получили бы вероятность единогласного осуждения, равную всего лишь 0,07.

3.7. Наше мнение (2002) о книге Курно весьма отрицательно.

4.1. Мгновенная замена повреждённого элемента означает, что всё рассуждение теряет смысл.

4.2. Строгое доказательство ЦПТ удалось только Ляпунову и Маркову. Ныне (Золотарёв 1999, с. 794) ЦПТ понимается шире, чем раньше, а именно как стремление к нормальному закону или к близкому к нему распределению.

4.3. Известно, что термин *нормальное распределение* ввёл Пирсон, чтобы не думать о сравнительных заслугах Гаусса и Лапласа. Гаусс пришёл к нормальному распределению в 1809 г. как к единственному закону ошибок, а в 1823 г. (окончательно) отказался от своего прежнего вывода. Понятие истинного значения измеряемой константы (см. ниже) восходит к Фурье и означает предел среднего арифметического из соответствующих наблюдений (Шейнин 2007).

4.4. Следовало бы добавить, что Муавр доказал свою формулу совсем иначе, и что выражение теорема *Муавра – Лапласа* восходит к Маркову. Лаплас ввёл поправку за счёт конечности числа наблюдений. Понятия равномерной сходимости в его время ещё не существовало.

4.5. Автор несколько раз употребляет выражение *чистая случайность*, которое следует понимать, как *только случайное*.

4.6. Об этой теории ничего не сказано, а *риск* усиленно изучается в страховом деле.

4.7. В 1766 г. Даниил Бернулли опубликовал своё знаменитое исследование оспы и вариоляции (*не вполне безопасной* профилактики оспы, которая постепенно исчезла после введения оспопрививания по Дженнеру). Инокуляция детей безусловно включала моральную проблему.

4.8. Даже ЦПТ Лаплас по существу не доказал. *Его исключительная интуиция позволила ему получать правильные выводы из нестрогих, а порой просто путаных рассуждений* (Гнеденко и Шейнин 1978, с. 105; Шейнин (2019, § 8.2-4).

4.9. Можно полагать, что обратное утверждение тоже верно, и таким образом устойчивость определяется постоянством параметров единого закона распределения ошибок наблюдения в течение рассматриваемого периода.

4.10. Теория ошибок есть приложение статистического метода (т. е. статистической теории, математической статистики) к обработке наблюдений. Медицинская и звёздная статистика это приложения того же метода к соответствующим наукам. Мы придерживаемся утверждения Пирсона (1892/1911, с. 15): *Единство всей* (данной) *науки состоит только в её методе*. Оно, кстати, отметает прежние представления о том, что статистика не наука, потому что не имеет собственного предмета.

Библиография публикатора

Бернштейн С. Н. (1926, франц.), Распространение предельной теоремы теории вероятностей на суммы зависимых величин. В книге автора *Собр. соч*., т. 4. М, 1964, с. 121 – 176.

Гнеденко Б. В., Шейнин О. Б. (1978), Теория вероятностей. Глава в книге *Математика XIX века*. М., с. 184 – 240. Редакторы А. Н. Колмогоров, А. П. Юшкевич.

Золотарёв В. М. (1999), Центральная предельная теорема. В книге Прохоров, редактор, с. 794 – 796.

Интеллектуальные (2016), *Интеллектуальные и информационные системы в медицине*. Москва – Берлин.

Прохоров Ю. В. редактор (1999), *Вероятность и математическая статистика. Энц*. М.

Шейнин О. Б., Sheynin O. (1998), Statistical thinking in the Bible and the Talmud. *Annals of Sci*., vol. 55, pp. 185 – 198.

--- (2002), Теоретико-вероятностное наследие А. О. Курно. *Историко-математические исследования*, вып. 7 (42), с. 301 – 316.

--- (2007), The true value of a measured constant and the theory of errors. *Hist. Scientiarum*, vol. 17, pp. 38 – 48.

--- (2019), *Теория вероятностей. Исторический очерк*. Берлин. **S, G,** 11.

Fienberg S. E. (1971), Randomisation and the 1970 draft lottery. *Science*, vol. 171, pp. 255 – 261.

Gelfand A. E., Solomon H. (1973), A study of Poisson's models for jury verdicts in criminal and civil trials. *J. Amer. Stat. Assoc.*, vol. 68, pp. 271 – 278.

Laplace P. S. (1812), *Théorie analytique des probabilités. Oeuvr. Compl.*, t. 7. Paris, 1886.

Pearson K. (1892, англ.), *Грамматика науки*. СПБ, 1911.

Poincaré H. (1896, франц.), *Теория вероятностей*. Ижевск, 1999.

II

В. Н. Тутубалин

Статистическая обработка рядов наблюдений

Москва, 1973

Введение

Как известно, *факты – воздух учёного*. В наш век точных научных методов слово *наблюдать* обычно означает *измерять*, и факты, с которыми мы имеем дело, как правило, выражаются в цифрах. В любом научном учреждении вам покажут длинные ряды цифр, которые также представлены в виде графиков, нарисованных разноцветными карандашами на миллиметровой бумаге. Всё это – ряды наблюдений. Спрашивается, какую пользу можно извлечь из такого разноцветного великолепия, на собирание которого тратятся долгие годы труда многих людей?

Часто бывает, что ряды наблюдений приводят к тому или иному очевидному выводу. Например, после введения в медицинскую практику антибиотиков резко снизилась смертность от большинства инфекционных болезней. Для того, чтобы сделать такой вывод, не требуется никакой математической обработки – данные говорят сами за себя. Но в других случаях выводы могут не быть столь бесспорными, и мы прибегаем к статистической обработке, стремясь повысить надёжность делаемых выводов за счёт применения математических методов.

Важно представлять себе, что во многих случаях статистическая обработка может быть полезна, но, пожалуй, ещё в большем числе случаев она бесполезна и иногда даже вредна, потому что толкает на ложные выводы. Так, бесполезно лечить антибиотиками болезни вирусного происхождения.

О тех случаях, в которых применение статистической обработки является научно обоснованным, и рассказывается в этой брошюре.

Глава 1. Две основные математические модели ряда наблюдений

1.1. Зачем нужна статистическая обработка? Как говорилось во **Введении**, статистическая обработка может быть и совсем не нужна. Чтобы привести ещё один пример такого рода, предположим, что мы стремимся повысить надёжность работы оборудования, и что нам удалось найти какое-то профилактическое мероприятие, которое явно снижает число отказов. Наш ряд наблюдений (допустим, количество отказов за разные годы) с несомненностью подтверждает эффективность мероприятия, и мы можем быть довольны. Но человеческая натура такова, что всё время хочет чего-то лучшего, и если отказов стало мало, то мы, конечно, захотим, чтобы их не было совсем. Пусть мы предлагаем новое мероприятие, и хотим подтвердить его эффективность, показав, что количество отказов за год снизилось ещё больше.

Можно ручаться, что это будет не так уж просто. Если отказов за год мало, то их число будет заметно подвержено влиянию случайных причин. Это ещё не означает, что его можно изучать чисто вероятностными методами, потому что для этого требуется статистическая однородность, которой может и не быть. О ней можно прочесть, например, в нашей брошюре [i].

Но вероятностная модель позволяет сделать некоторые важные выводы, которые необходимо иметь в виду. Кроме того, в условиях, когда хороший технический результат уже достигнут, а мы добиваемся лучшего, сравнительно часто возникает статистическая однородность. Рассмотрим поэтому вопрос о проверке эффективности нового мероприятия, предполагая, что вероятностная модель для числа отказов подходящая.

Если признать вероятностный подход вообще, то для редких событий весьма естественно признать и распределение Пуассона. Это означает, что для числа отказов в год справедлива формула

$$P(\mu = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Здесь λ – параметр, равный среднему числу отказов в год. Допустим, что с начала некоторого года мы ввели новое мероприятие, и в течение этого года не произошло ни одного отказа, тогда как среднее число отказов за прошлые годы равно 2. Можно ли считать, что нововведение эффективно?

Число 2 получено по статистическим данным за прошлые годы. Оно не обязано точно совпадать с нынешним значением λ, но пока мы предположим, что разностью можно пренебречь. Итак, $\lambda = 2$. Но в таком случае вероятность того, что чисто случайно отказов не будет, т. е. что $\mu = 0$, равна

$$P(\mu = 0) = e^{-2} \approx 1/7.$$

Поэтому, если на основании имеющихся данных мы признаём нововведение эффективным и выплатим авторам премию, то с вероятностью 1/7 выплата окажется напрасной. Итак, 1/7 всех научных сотрудников, предложивших нечто бесполезное (например, окуривание оборудования благовониями), получит премию. Беда, однако, не столько в том, что будут зря потрачены премиальные деньги, сколько в том, что получат признание совершенно ложные воззрения, ведь выплата премии означает признание. И вот окуривание оборудования благовониями уже введено в информационные системы (которые не разбираются, что вздор, а что нет) и через них во всеобщую практику. На следующий год 1/7 тех, кто воспользовался этим *методом*, вновь счастлива и публикует восторженные статьи, а 6/7, которые не добились успеха, молчат, потому что из неудачи статью не напишешь. Процесс нарастает как лавина, в вузах уже создаются кафедры благовоний, проводятся конференции, пишутся диссертации, учебники и т. д.

Нарисованная картина очень печальна, но не является беспредметной. Примеры такого рода имеются. В брошюре [i] рассказывается о присяжных Пуассона. Ниже мы также приведём подобные примеры. Размышления над нарисованной картиной *науки о благовониях* заставляет, как нам кажется, по- новому оценить достоинства настоящей науки, этого удивительного достижения коллективного разума. *Науки о благовониях* время от времени, и даже часто возникают, процветают, ведь чем больше людей ими занимается, тем больше сообщений об успехах (в нашей модели успеха добивается каждый год 1/7 всех занимающихся), но не живут долго. Всегда находится человек, который наносит им решительный удар. Остаются лишь настоящие ценности.

В этом процессе самоочищения науки далеко не последняя роль принадлежит вероятностным методам, хотя, конечно, здесь нелепо говорить об их монополии. Надо уметь ответить на вопрос: а не мог ли наблюдаемый эффект получиться чисто случайно? Но математическая статистика, как всякая наука, может иметь свои разделы *о благовониях*. Мы сейчас рассмотрим общую структуру статистических методов: что в них достоверно, а что условно, и на каких предпосылках они основаны.

1.2. Роль математических моделей. Для любой статистической обработки нужно сначала предложить математическую модель явления, т. е. сказать, какие величины случайны и неслучайны, какие зависимы и независимы и т. д. Иногда можно встретить заблуждение, согласно которому если уж некоторая величина не является детерминированной, т. е. если мы не можем заранее точно предсказать её значения, то её можно рассматривать как случайную. Это совершенно неверно, потому что случайность требует статистической устойчивости. Итак, неопределённость не есть, вообще говоря, случайность. Или, если угодно, кроме детерминированных и случайных величин, бывают ещё неопределённые, с которыми неизвестно что делать.

Математическая модель может содержать либо детерминированные, либо случайные величины, либо те и другие вместе, но, при нынешнем состоянии науки, не может содержать неопределённых. Следовательно, искусство выбора математической модели состоит в том, чтобы те неопределённые величины, с которыми всегда приходится иметь дело на практике, приближённо представить либо как детерминированные, либо как случайные. При этом требуется ещё, чтобы значения детерминированных величин или распределения вероятностей случайных величин можно было бы определить из того экспериментального материала, которыми мы располагаем (или в принципе можем располагать).

Вернёмся к задаче об определении эффективности нового профилактического мероприятия. Мы имеем ряд наблюдений

$\mu_1, \mu_2, \ldots, \mu_n, \mu.$ (1.2.1)

Здесь указаны количества отказов оборудования за годы, предшествовавшие введению мероприятия, а μ – число отказов за тот год, когда испытывалось нововведение. Какова здесь математическая модель? В случае редких отказов довольно разумно считать, что ряд наблюдений (1.2.1) составлен из случайных величин. Однако, говоря *случайная величина*, мы обязываем себя указать статистический ансамбль экспериментов, в которых эта величина реализуется. Можно пойти двумя путями: либо считать, что числа отказов за разные годы до введения нового мероприятия являются реализациями случайной величины, либо представить себе результаты эксплуатации многих комплектов оборудования, идентичных нашему, в идентичных условиях. В первом случае указанные числа обязаны быть одинаково распределёнными, во втором это не обязательно. Иначе говоря, принимая распределение Пуассона, получаем в первом случае

$$E\mu_1 = E\mu_2 = \ldots = E\mu_n = \lambda,$$

а во втором случае можно положить

$$E\mu_1 = \lambda_1, E\mu_2 = \lambda_2, \ldots, E\mu_n = \lambda_n,$$

где числа λ_i могут быть различны.

Теоретически второй случай более общий и потому на первый взгляд более привлекателен, чем первый. Но мы сейчас увидим, что этот путь не ведёт ни к чему, и потому должен быть оставлен.

Действительно, нам необходимо знать, каков был бы пуассоновский параметр $\lambda = E\mu$ для числа отказов в год когда испытывается нововведение при его неэффективности. Но если между числами λ_i нет никаких связей, то λ никак не связано с нашими наблюдениями μ_i. Итак, мы не в состоянии узнать λ. Далее, в качестве оценок $\hat{\lambda}_i$ параметров λ_i мы должны взять оценки, построенные по единственной реализации (если у нас есть наблюдения над единственным комплектом оборудования), т. е. нам не остаётся ничего другого, как положить $\hat{\lambda}_i = \mu_i$, но такие оценки очень грубы.

Итак, выбрав очень общую модель, мы не в состоянии определить её параметры. Так бывает всегда. С другой стороны, частная модель позволяет дать лучшую оценку

$$\hat{\lambda} = \overline{\mu}. \qquad (1.2.2)$$

В этом случае естественно считать, что при неэффективности нововведения

$$\lambda = E\mu \approx \hat{\lambda} = \overline{\mu}. \qquad (1.2.3)$$

Таким образом, частная модель, вообще говоря, позволяет решать нашу задачу. Но у этой модели другой недостаток, она может быть неверна. Например, из-за старения оборудования среднее число аварий может увеличиваться с каждым годом:

$$E\mu_1 < E\mu_2 < ... < E\mu_n < E\mu.$$

Если в этом случае принять оценку (1.2.3), то занизим фактическое значение Eμ. Если $\hat{\lambda} = 2$, а фактически Eμ= 4, то на самом деле

$$P\{\mu = 0\} = e^{-4} \approx 1/55,$$

мы же будем думать, что $P \approx 1/7$, § 1.1, и не признаем нововведения эффективным, тогда как оно почти наверняка эффективно.

Итак, при построении статистической модели приходится выбирать между Сциллой и Харибдой, т. е. между общей моделью, которая бесполезна, потому что нельзя определить её параметров, и частной, которая может быть неверна, а потому приведёт к неправильным выводам. Неизвестно только, что именно Сцилла, а что Харибда.

Допустим, что мы приняли частную модель, т. е. объявили величины μ_i одинаково распределёнными случайными величинами. Кончаются ли на этом все необходимые предположения? Нет, не кончаются, потому что нам крайне необходимо знать, как велика может быть ошибка приближённого равенства (1.2.3). Например, если $\hat{\lambda}=2$, может ли настоящее значение λ равняться 4? Иначе говоря, нужно уметь вычислять дисперсию величины (1.2.2). Она равна

$$D\hat{\lambda} = \frac{1}{n^2}[\sum_{i=1}^{n} D\mu_i + \sum_{i \neq j} \operatorname{cov}(\mu_i \mu_j)].$$

Для закона Пуассона

$$D\mu_i = E\mu_i = \lambda,$$

а для грубой оценки можно принять, что $D\mu_i \approx \hat{\lambda} = \overline{\mu}$.

Но о ковариациях мы ничего не можем сказать. Имеющихся данных обычно далеко недостаточно, чтобы оценить их. Не остаётся ничего лучшего, чем предположить величины μ_i независимыми, т. е. считать, что ковариации равны нулю.

Итак, мы приходим к модели независимых одинаково распределённых случайных величин, т. е. к выборке. Читатель вероятно согласится, что наши рассуждения, если и не доказывают логически, что лишь модель выборки может быть полезна, всё же достаточно убедительны, чтобы показать, что из круга идей, ведущего к модели выборки, вырваться трудно.

Поэтому модель выборки очень популярна, с ней стараются работать, если только нет доказательств, что она неверна.

Главы математической статистики, посвящённые выборкам, бесспорно составляют лучшую и наиболее разработанную часть этой науки. Но модель выборки достаточно (и даже слишком) часто бывает неверна. Мы видели, что наблюдения μ_i не образуют выборки, если оборудование стареет. То же самое верно, если есть приработочный период, т. е. если в начале эксплуатации устраняются производственные дефекты, после чего число отказов падает. Есть и другие способы нарушить идентичность распределения величин μ_i.

Возможно и нарушение их независимости. Если, например, отказ оборудования влечёт за собой основательный ремонт с широкой заменой изношенных, но ещё способных работать деталей, то возникает отрицательная корреляция между μ_i и μ_j. Если же отказ одной детали усиливает износ других, а замены не происходит, то возникает положительная корреляция.

Понятно, что, вводя модели, отличающиеся от модели выборки, мы должны характеризовать их отличие от выборки небольшим числом параметров, которые можно определить либо теоретически, либо по имеющимся статистическим данным. Сложные модели, как уже говорилось, абсолютно бесполезны. Практически можно учесть либо отклонения от идентичности распределений, задаваемые детерминированными функциями, либо отклонения от независимости при сохранении идентичных распределений. Мы сейчас перейдём к соответствующим моделям.

Надо иметь в виду, что эти модели, как и модель выборки, достаточно условны. Правда, если модель правильна, то выводы, которые мы делаем, получаются уже чисто математическими средствами и потому достоверны. Но в общем всё зависит от модели. Статистические методы имеют такую же, не бОльшую, но и не меньшую достоверность, как выводы других наук, которые пользуются математическими средствами, как физика, астрономия или сопротивление материалов. В практических вопросах эти науки могут дать нам путеводную нить, но никаких особых гарантий того, что мы правильно воспользовались ими.

1.3. Модель тренда с ошибкой. В математической модели ряда наблюдений всегда что-то детерминировано, а что-то случайно. Мы рассмотрим модель, в которой ряд наблюдений

$$x_1, x_2, \ldots, x_n$$

задаётся формулой

$$x_i = f(t_i) + \delta_i. \qquad (1.3.1)$$

Здесь t_i это значение некоторой детерминированной переменной, которая характеризует i-й эксперимент, $f(t)$, некоторая детерминированная функция (тренд), и δ_i, случайная величина, обычно называемая ошибкой этого эксперимента.

Речь идёт, таким образом, о том[1.1], что *господь бог* определил истинную зависимость в виде $f(t)$, так что в i-том эксперименте мы должны были бы наблюдать $f(t_i)$, но *дьявол* вносит в наши наблюдения ошибку δ_i.

Например, $f(t)$ может изображать ту или иную координату космического объекта в зависимости от времени t, и тогда x_i будет нашим измерением этой координаты в момент t_i. Конечно, вмешательство *дьявола* δ_i может носить детерминированный, случайный или вообще неопределённый характер. Например, у наблюдений x_i может быть систематическая ошибка, т. е. ожидание $E\delta_i$ не обязательно равно нулю. Можно допустить, что $E\delta_i = C$ и не зависит от i, но можно считать, что $E\delta_i = \varphi(t_i)$ является функцией от t_i. Ещё хуже получится, если $E\delta_i$ зависит от не контролируемой в опыте переменной u_i. Ни в каких подобных случаях статистическая обработка не может избавить нас от ошибок.

Но если наблюдения организованы достаточно тщательно, то можно надеяться, что ошибки будут только случайными в том смысле, что соблюдается статистическая однородность и отсутствует систематическое смещение: $E\delta_i = 0$. Точнее говоря, систематическая ошибка будет достаточно мала, так что ей можно пренебречь.

Такие ситуации и составляют область применения статистических методов. После того, что было сказано в § 1.2, ясно, что на ошибки δ_i надо наложить простейшие статистические предположения. Чаще всего их предполагают независимыми и одинаково распределёнными. Обычно добавляют предположение нормальности.

Лишь одно отклонение от модели выборки для δ_i вошло в употребление. Иногда считают, что дисперсии δ_i не одинаковы, а пропорциональны как-то заданным числам. Иначе говоря, считают, что известны числа w_i, которые называются весами наблюдений и для которых имеют место равенства $w_1 D\delta_1 = w_2 D\delta_2 = ... = w_n D\delta_n = \sigma^2$. В таком случае дисперсии обратно пропорциональны весам

$$D\delta_i = \frac{\sigma^2}{w_i}.$$

Мы описали предположения, накладываемые на случайные составляющие наших наблюдений. Перейдём теперь к описанию детерминированной составляющей $f(t)$, которая называется также *трендом*. Простейший и классический случай состоит в том, что функция $f(t)$ имеет вполне определённый вид, но зависит от некоторых неизвестных параметров

$c_1, c_2, ..., c_k$:

$$f(t) = F(t, c_1, c_2, ..., c_k). \qquad (1.3.2)$$

Здесь функция F задана известной нам формулой или алгоритмом вычислений. Например, при движении космического объекта под параметрами можно понимать его координаты и скорости в любой определённый момент времени. Можно ввести и другие, более удобные параметры. Тогда, исходя из законов движения Ньютона (если объект не имеет двигателя), любую координату $f(t)$ однозначно определяют по параметрам.

Задача состоит в том, чтобы по известным наблюдениям (1.3.1) определить оценки $\hat{c}_i$ параметров. Решение задачи известно, это применение метода наименьших квадратов Гаусса. Оно состоит в том, что $(\hat{c}_1,...,\hat{c}_k)$ следует выбрать таким образом, чтобы при известных наблюдениях в этой точке достигался минимум функции

$$\sum_{i=1}^{n}[x_i - F(t_i; c_1,...,c_k)]^2 .$$

Чаще встречается случай, когда истинная зависимость $f(t)$ неизвестна, но и тогда пользуются зависимостью (1.3.2), но функцию F выбирают более или менее произвольно, например, принимают многочлен и вновь применяют метод наименьших квадратов.

Такая неклассическая ситуация, когда $f(t)$ заранее не известна, требует более подробного анализа, см. конкретный пример в следующей главе. Сейчас мы, однако, опишем совершенно иную модель, также применяемую для статистической обработки рядов наблюдений.

1.4. Модель случайного процесса. В только что рассмотренной модели тренда с ошибкой основное внимание обращается на выделение детерминированной составляющей, тренда $f(t)$. Сами значения x_i вида (1.3.1) не являются случайными величинами, случайны лишь добавки δ_i, которые считаются ошибками, помехами и вообще *кознями дьявола*.

Возможен иной подход, когда случайность рассматривается как основное свойство наблюдаемого ряда, который мы теперь обозначим через

$$\xi_1, \xi_2, ..., \xi_n. \tag{1.4.1}$$

Простейшая модель здесь такова: весь этот набор считается реализацией n-мерной случайной величины. Если есть возможность многократного повторения эксперимента, который обеспечивает подобную реализацию, т. е. получение многих рядов наблюдений в аналогичных, статистически однородных условиях, то такая модель может оказаться работоспособной. Но чаще мы имеем только один ряд наблюдений, и по одной реализации n-мерной случайной величины, конечно же, нельзя восстановить никаких распределений вероятностей, так что модель n-мерного распределения абсолютно бесполезна.

Но если предположить, например, что совместное распределение величин ξ_1, ξ_2, таково же, как распределение величин ξ_2, ξ_3; ξ_3, ξ_4, и т. д., то пары $(\xi_1, \xi_2), (\xi_2, \xi_3),..., (\xi_{n-1}, \xi_n)$ дают много реализаций этого двумерного распределения, хотя быть может не независимых друг от друга. Таким образом, в принципе двумерное распределение можно определить.

Удобно несколько расширить математическую модель. Рассмотрим бесконечную в обе стороны последовательность

$$... \xi_{-1}, \xi_0, \xi_1, \xi_2, ..., \xi_n, \xi_{n+1}, ... \qquad (1.4.2)$$

случайных величин. Такая последовательность называется *случайным процессом*. Мы считаем, что теоретически существуют распределения вероятностей любого конечного набора

$$(\xi_{k_1}, \xi_{k_2}, ..., \xi_{k_m}) \qquad (1.4.3)$$

случайных величин. Наш ряд наблюдений (1.4.1) составляет часть бесконечной последовательности (1.4.2). По этим наблюдениям можно сделать какие-то заключения о всём этом случайном процессе (1.4.2) только в том случае, если в модель включено правило, позволяющее выражать распределения величин (1.4.3) при отрицательных и больших положительных значениях индексов через распределение наблюдаемых величин (1.4.1). Без такого правила модель случайного процесса бесполезна. В простейшем и наиболее естественном случае накладывается условие стационарности: распределение величин

$$(\xi_{k_1} + \tau, \xi_{k_2} + \tau, ..., \xi_{k_m} + \tau)$$

при любых τ совпадает с распределением (1.4.3). И модель случайного процесса состоит в том, что мы считаем наблюдения (1.4.1) частью реализации (1.4.2) стационарного случайного процесса.

Работая с моделью случайного процесса, обычно не идут дальше двумерных распределений и, кроме того, ограничиваются изучением корреляций между различными значениями случайного процесса. Несмотря на всю популярность понятия случайного процесса, можно привести совсем немного случаев, когда бы эта концепция позволяла должным образом описывать статистические свойства рядов наблюдений. Большинство работ по теории случайных процессов начинается словами *пусть дан случайный процесс с такими-то и такими-то характеристиками*. Но совсем мало работ, в которых характеристики случайных процессов в самом деле определяются теоретически или экспериментально. В таком виде теория случайных процессов подходит для решения модельных задач: что произойдёт, если на некоторую систему воздействует белый шум заданной интенсивности? Но модельные задачи имеют лишь косвенное отношению к реальному поведению системы, потому что в реальных условиях на систему воздействует скорее всего не белый шум, и может быть даже не случайный процесс (нет статистической однородности). А между тем часто никто не изучает, что в действительности воздействует на систему, потому что подобное исследование сложно, трудно и дорого и гораздо легче ограничиться произвольными априорными предположениями.

Поэтому представляет интерес посмотреть, что получалось у наиболее выдающихся статистиков, когда они пытались изучать фактические данные с помощью модели случайного процесса. Довольно часто дело кончалось неудачей. Мы расскажем об этом в третьей главе. Кратко упомянем также о статистической теории турбулентности, в которой понятие случайного процесса было применено с блестящим успехом.

Глава 2. Метод наименьших квадратов

Этот метод придумал и ввёл во всеобщее употребление Гаусс[2.1]. Классический случай, который он рассмотрел, заключался в том, что между членами ряда наблюдений

$$x_1, x_2, \ldots, x_n$$

должны были бы соблюдаться некоторые известные соотношения, не будь в них ошибок. Например, в случае траектории космического тела все члены ряда наблюдений можно было бы выразить через несколько первых членов, если бы они были известны совершенно точно.

Этот классический случай сравнительно легко можно исследовать в рамках математической статистики. Но при практической реализации метода наименьших квадратов (МНКв) могут возникнуть более или менее значительные вычислительные трудности, однако мы не будем их касаться. Другие трудности связаны с тем, что могут не выполняться предположения модели тренда с ошибкой. Например, при радиолокационном измерении дальности ошибки последовательных определений, по-видимому, нельзя считать независимыми случайными величинами. Вообще неясно, имеют ли эти ошибки статистический характер, так что статистические методы здесь ненадёжны или вообще бессильны. Но зато сами наблюдения имеют высокую точность и могут быть сделаны в большом числе, а потому статистические методы здесь и не нужны.

При всех достоинствах классического случая он имеет тот недостаток, что сравнительно редко встречается. Гораздо чаще мы убеждены в том, что наши наблюдения могут быть аппроксимированы гладкой зависимостью

$$x_i \approx f(t_i),$$

где t_i – переменная, описывающая условия i-го эксперимента. Но мы не имеем точного вида функции $f(t)$, так что здесь применяются приёмы, очень сходные с приёмами классического случая, но, как показывает их анализ, не имеющие какого-либо математического обоснования. Математическая статистика широко использует математику, но не сводится к этой сравнительно очень ясной науке. Статистика это скорее ремесло, и как таковое имеет свои тайны, с которых мы и начнём[2.2].

2.1. Тайны статистического ремесла. Желая применять МНКв, мы в большинстве случаев можем воспользоваться написанными раз навсегда программами для ЭВМ. Достаточно ввести в машину данные наблюдений, дождаться, пока будут сделаны вычисления, и из печатающего устройства выползет бумажная лента с формулой для сглаживающей кривой. Но неправ будет тот, кто целиком поручит машине всю эту процедуру. Совершенно необходимо представить имеющиеся данные в обозримом виде и хотя бы взглянуть на чертёж[2.3].

Человеческий глаз в состоянии заметить в материале наблюдений такие особенности, которые недоступны машине. Например, если первая половина наблюдений лежит выше сглаживающей кривой, а вторая ниже, то, очевидно, нарушается предположение о независимости ошибок в модели тренда с ошибкой. В таком случае никакие машинные вычисления не имеют никакого смысла, но машина сама по себе не в состоянии заметить подобных особенностей числового материала. Конечно, можно запрограммировать соответствующую проверку, но беда в том, что мыслимых особенностей в данных слишком много, чтобы можно было их все включить в программу исследований.

Применение любого заданного статистического критерия естественно поручить машине, но машина почти бессильна сформулировать те критерии, которыми нужно испытывать имеющийся материал. Формулировать критерии должен статистик, и делает он это на основании глазомерной оценки статистических данных, которые поэтому должны быть представлены в наглядном виде. Следовательно, статистическое ремесло основано в первую очередь на глазомерной оценке.

Тот, кто полностью полагается на автоматические машинные вычисления, лишается возможности проверять статистическую модель. В результате, чем более автоматизируется обработка, тем меньше доверия она заслуживает. Но если статистический материал требует глазомерной оценки, то это означает, что с функциями одной переменной мы умеем работать хорошо (потому что их можно изобразить на графике), с функциями двух переменных много хуже (их можно изображать изолиниями подобно рельефу на топографических картах), а с функциями большего числа переменных совсем не умеем работать[2.4]. Это та область, в которой мы можем рассчитывать лишь на помощь вычислительных машин. Здесь сделаны первые шаги, появились такие направления, как многомерный статистический анализ и планирование экстремальных экспериментов, но до создания действительно надёжных методов ещё очень далеко. Приёмы только что упомянутых направлений иногда работают, иногда нет, а почему – мы не знаем. Главный недостаток здесь – низкий моральный уровень исследований, обыкновение выдавать желаемое за действительность. В результате неизвестно, чему можно верить, а чему нет.

Итак, желая применить МНКв, нужно прежде всего построить график ряда наблюдений. Читатель составит себе представление о том, что при этом получается, если взглянет на ломаную линию рис. 1 [ни один рисунок автора не воспроизведён]; пока неважно, что она означает. Очевидно, что *ломаная такого рода колеблется около некоторой плавно меняющейся* (?) *кривой. Эта кривая и выражает истинную закономерность, а колебания ломаной вызваны случайными причинами, и не имеют отношения к истинной закономерности.*

Выделенные фразы обычно составляют всю информацию, которую мы имеем об изучаемой истинной зависимости. Понятно, что подобная информация слишком расплывчата и неопределённа, чтобы непосредственно допускать какое-то научное использование. Вывод из глазомерного исследования необходимо превратить в математическую модель, которую можно было бы использовать при статистической обработке. Переход от глазомерной оценки к математической модели – вторая тайна статистического ремесла.

В случае МНКв чаще всего применяется модель вида

$$x_i = P(t_i) + \delta_i,\ i = 1, 2, \ldots, n.$$

Здесь $P(t) = a_0 + a_1t + \ldots + a_mt_m$ – многочлен, коэффициенты которого оцениваются по МНКв[2.5]. Какие имеются доводы в пользу такой модели, которая приводит к процедуре так называемого *сглаживания многочленом*? Обычно говорится, что если на самом деле верна другая модель

$$x_i = f(t_i) + \delta_i,\ i = 1, 2, \ldots, n,$$

где функция $f(t)$ отлична от многочлена, то можно воспользоваться теоремой Вейерштрасса, в силу которой функцию $f(t)$ можно приблизить многочленом $P(t)$ с любой точностью. Но ссылка на эту теорему здесь неуместна, потому что в соответствии с ней любая непрерывная функция может быть приближена многочленом достаточно высокой степени. На практике, однако, степень m многочлена $P(t)$ стараются выбирать не высокой, а низкой. Действительно, если n – число экспериментальных точек, то многочлен степени $(n – 1)$ пройдёт через все экспериментальные точки $x_1, \ldots, x_n$, и тем самым включит случайные погрешности δ_i. Итак, многочлен более высокой степени окажется не ближе к истине, а дальше, чем многочлен более низкой степени.

Далее, теорема Вейерштрасса действует и для функций нескольких переменных. Но если считать типичной функцией двух переменных земной рельеф, изображённый горизонталями на топографической карте, и попытаться приблизить его многочленом, то из этой попытки обычно ничего не выходит: степень многочлена должна быть слишком высока. Итак, одно дело – теоретическая теорема Вейерштрасса, другое дело – практическое сглаживание.

Следовательно, сглаживание многочленом не имеет каких-либо математических оснований, и успех этой процедуры составляет ещё одну тайну статистического ремесла. Как можно объяснить этот довольно часто встречающийся практический успех? По-видимому, человеческий глаз хорошо чувствует ход графиков аналитических функций, в частности, многочленов. Вспомним, например, каким образом студентов учат рисовать графики на упражнениях по математическому анализу. Требуется вычислить совсем немного точек – разрывы, экстремумы, иногда и точки перегиба. По этим типичным точкам функция восстанавливается совсем неплохо. Можно предположить, что, увидев ломаную линию, человек в состоянии уловить, хорошо ли приближается многочленом истинная зависимость. Статистическая процедура сглаживания многочленом вероятно применяется только в тех случаях, когда рассматривание графика создаёт надежду на успех.

Но положение сразу меняется, если попытаться автоматизировать всю процедуру. В этом случае машинным расчётам будут подвергаться любые данные без их предварительной оценки, и процент успешных сглаживаний резко сократится. Совсем сложна задача сглаживания в случае функции многих переменных, когда мы не можем изобразить ни экспериментальные данные, ни результаты сглаживания. Мы даже не можем оценить, достигнут ли успех или нет.

2.2. Практический пример сглаживания многочленом. Пора объяснить, откуда взялись наблюдения, представленные на рис.1; повторим: он не воспроизведён. Во Всесоюзном научно-исследовательском институте электроэнергетики (ВНИИЭ) изучались повреждения статорной изоляции крупных турбогенераторов. Результаты опубликованы в работах [1], и рис. 1 взят из них.

Общее число аварий складывается из аварий отдельных генераторов. Довольно рано было решено, что вероятность аварии данного генератора пропорциональна общей площади его изоляции и мало зависит от прочих конструктивных и эксплуатационных данных. Заметим, что тогда почти не было генераторов с водородным охлаждением. Поэтому было решено рассматривать поведение единицы площади изоляции, отвлекаясь от всех прочих особенностей.

Был поставлен вопрос о старении изоляции, т. е. о возрастании вероятности аварии с увеличением возраста изоляции. Значения частот аварий образуют ломаную и этот разброс возрастает от малых значений t к большим, что связано с резким уменьшением площади изоляции, т. е. с сокращением экспериментального материала.

Нас интересуют значения вероятностей $p(t)$, равные вероятности аварии единицы площади изоляции возраста t при её эксплуатации в течение единицы времени, 10^4 часов, что соответствует примерно полутора календарным годам. Хотя объём экспериментального материала при небольших t велик, сами вероятности $p(t)$ малы, составляя 0,01 – 0,02, так что их прямое определение по частотам даёт очень большую ошибку.

Как известно, среднее квадратическое уклонение частоты μ_i/S_i, где μ_i это число аварий в интервале от $(i-1)$-й и i-й единиц времени, а S_i, соответствующая площадь, даётся формулой

$$\sqrt{\frac{p(t_i)[1-p(t_i)]}{S_i}}.$$

Здесь $t_i = 10^4 i$ часов. Например, для $t_1 = 10^5$ имеем $p(t_i) \approx 0{,}02$, $S_i = 200$, так что среднее квадратическое уклонение примерно равно 0,01 или 50% самой величины $p(t)$.

Итак, естественно попытаться повысить точность определения величин $p(t_i)$ за счёт сглаживания, потому что при этом в оценку входят все экспериментальные данные. Сглаживание соединяет все данные в одной обработке и потому повышает точность. Но в таком случае необходима статистическая модель. Наблюдаемые количества аварий (см. 1.2.1) естественно считать случайными величинами с пуассоновским распределением вероятностей. Понятно, что

$$\mathrm{E}\mu_i = S_i p(t_i).$$

Несколько труднее решиться на то, чтобы объявить величины μ_i независимыми. Но здесь помогают следующие соображения, которые применимы к любым редким событиям. Рассмотрим, например, μ_1, μ_2. Конечно, аварии, случившиеся в течение первого интервала времени, влияют на поведение изоляции во втором интервале, но их влияние ограничивается теми машинами, на которых они случились. Поскольку доля таких машин невелика, влияние тоже незначительно.

Признание независимости величин μ_i полностью задаёт математическую модель. Правда, эта модель связана не с наиболее удобным нормальным, а с пуассоновским распределением. Далее, дисперсии

$$D\mu_i = \mathrm{E}\mu_i = S_i p(t_i)$$

зависят от $p(t_i)$, которые мы и хотим определить. Здесь помогает переход к величинам

$$v_i = 2\sqrt{\mu_i}, \qquad (2.2.1)$$

который сильно выравнивает дисперсии.

Величины v_1, v_2, ..., v_n, от которых мы вернёмся к (1.2.1), сглаживаются. Эта процедура сама по себе несложна, но использует некоторые, подробно описанные [1] математические ухищрения. Окончательным ответом следует считать приближённое выражение для $p(t)$

$$p(t_i) = p(x_i) \approx (1/4)[b_0 - 0{,}1333b_2 + b_2x^2]^2 + 0{,}35/S_i,\ x_i = t_i/22,$$

где t_i измерено в принятых единицах времени, а последний член, как оказывается, надо прибавить для учёта систематической погрешности, которая возникает при переходе к v_i (2.2.1).

Для b_0 и b_2 и их дисперсий указаны оценки

$$\hat{b}_0 = 0{,}225,\ \hat{b}_2 = 0{,}20,\ D\hat{b}_0 = 2{,}12\cdot 10^{-4},\ D\hat{b}_2 = 44{,}5\cdot 10^{-4}.$$

Величины $0{,}35/S_i$ сглажены некоторым многочленом.

Статистически аккуратная работа с вероятностями аварий должна использовать наш ответ именно в таком виде. Но, поскольку он крайне мало нагляден, мы представили его в другой, упрощённой форме. С этой целью для (b_0, b_2) был указан 80% доверительный прямоугольник и построены три кривые, $p_1(t)$, $p_2(t)$ и $p_3(t)$. Кривая $p_2(t)$ даёт наилучшую оценку истинной вероятности аварии, которую мы можем предложить. Она получается, если вместо b_0 и b_2 подставить их оценки, см. выше. Эту кривую рекомендуется использовать для расчётов.

Другие две кривые получаются, если истинную точку (b_0, b_2) заменить соответственно левой нижней и правой верхней вершинами доверительного прямоугольника. Они дают представление о порядке возможной ошибки кривой $p_2(t)$, но, строго говоря, не являются границами доверительной области для истинной кривой. Доверительную область для $p(t)$ можно строить разными способами, хотя, опять же, строго говоря, она не нужна, потому что вся информация, употребляемая для её построения, суммирована в указанных дисперсиях $D\hat{b}_0$ и $D\hat{b}_2$. Можно смотреть на область между $p_1(t)$ и $p_3(t)$ как на некоторый приближённый (но верный по своему порядку) вариант доверительной области.

Принципиальное значение имеют различные варианты контроля использованной нами модели с применением статистических критериев. Никакая проверка в статистике не является исчерпывающей, но ряд хорошо вышедших проверок даёт тем не менее чувство уверенности в результатах. Мы подробно остановимся на всех этих проверках.

Простейший контроль – по окончательным результатам. Посмотрим внимательно на кривую $p_2(t)$, которая предлагается в качестве искомой зависимости $p(t)$. Не слишком ли велико отклонение фактических данных от этой кривой? Наибольшее отклонение достигается при $t = 20 \times 10^4$ часов. Величина $\mu_{20} = 4$, тогда как, если верить значению $p_2(20) = 0{,}041$, должно быть $\lambda_{20} = E\mu_{20} = p_2(20)$. $S_{20} = 0{,}041 \times 46 = 1{,}88$. Таким образом, в соответствии с законом Пуассона $P\{\mu_{20} \geq 4\} \approx 0.12$.

Итак, значение $\mu_{20} = 4$ само по себе не особенно значимо. Если же учесть, что это наибольшее из 22 событий, и что для одного из них вероятность 1/22 ≈ 0,05, вполне приемлема, то данное значение μ_{20} ни о каком расхождении $p_2(t)$ с экспериментальными данными не свидетельствует.

Можно составить выражение, сходное с суммой квадратов отклонений экспериментальных данных от гладкой кривой $p_2(t)$. Но лучше иметь дело не с величинами μ_i, а с v_i (2.2.1), потому что их дисперсия мало зависит от неизвестных вероятностей $p(t_i)$ и грубо говоря, равна 1.

Таким образом, мы имеем здесь случай, когда дисперсия наблюдений почти точно известна. Это связано с тем, что пуассоновское распределение зависит от одного параметра, а не от двух, как нормальный закон. В общем, и критерий с суммой квадратов уклонений даёт хорошее согласие с кривой $p_2(t)$.

Другая группа проверок связана с выбором степени и числа членов многочлена, аппроксимирующего истинную зависимость. Здесь мы также работаем с величинами v_i (2.2.1) и проверяем, что получится, если аппроксимировать эти величины различными многочленами степени до третьей включительно. Ясно, что искомый многочлен содержит член нулевой степени. Попробуем прибавить по очереди члены первой, второй и третьей степени. Наилучшее улучшение аппроксимации достигается при добавлении члена второй степени, т. е. при многочлене вида

$$c_0 + c_2 t^2. \qquad (2.2.2)$$

Теперь наступает очередь статистических проверок, которые показывают, что прибавление к этому многочлену членов первой и третьей степеней не улучшает аппроксимацию значимо. Подробности вычислений см. в сборнике [1]. В результате всех этих проверок мы получаем уверенность, что с помощью многочлена (2.2.2) мы в самом деле извлекли всю детерминированную составляющую, которую можно было извлечь из имеющихся данных.

Но при всём благополучии проверки гипотез, связанных со сглаживанием, мы абсолютно не проверяем основную гипотезу, что вероятность аварии единицы изоляции не зависит от конструктивных и эксплуатационных особенностей той машины, на которой эта изоляция установлена. Действительно, мы проверяем лишь пуассоновское распределение (и отчасти независимость) величин μ_i. Но это распределение возникнет и при неравных вероятностях аварий на различных единицах площади изоляции, лишь бы эти вероятности были достаточно малы.

Итак, важнейшая гипотеза статистической однородности различных единиц площади изоляции пока остаётся без проверки, и вообще по сводным данным не может быть проверена. Между тем основной технический интерес представляет именно выделение и изучение высокоаварийных и низкоаварийных машин, либо подтверждение того, что все машины имеют одинаковую аварийность. Сейчас мы посмотрим, каким образом можно решить эти вопросы.

2.3. Проверка статистической однородности. Первейшее условие для получения статистически однородной совокупности, и, так сказать, первейшая тайна статистического ремесла, состоит во внимательном отборе материала для исследования. Так, в данные рис. 1 не вошли аварии изоляции, вызванные посторонними для неё причинами, например, повреждением изоляции посторонними предметами, случайным попаданием высокого напряжения, перегрузками при коротком замыкании и т. д. Мы полагали, что небрежность слесаря, который оставил на изоляции обломок напильника, и другие подобные явления, хотя и называются *случайными* в обычном смысле, не являются таковыми в смысле теории вероятностей, так как не обладают статистической устойчивостью.

Отбор материала для обработки облегчался тем, что авария крупной машины есть из ряда вон выходящее событие, так что её причины тщательно исследуются и записываются в аварийный акт. Наиболее приемлемой для включения аварии в статистическую обработку оказалась формулировка: *причиной аварии является местный дефект изоляции*. Но в общем были включены все аварии, в актах о которых не было ясных указаний на повреждение изоляции посторонней причиной. При отборе материала статистик всегда должен держаться какого-то раз навсегда сформулированного принципа. Грубо говоря, аварии, включённые в статистическую обработку, составляют около половины общего числа аварий изоляции.

Итак, лишь около половины всех аварий, которые происходят на самом деле, возможно могут быть описаны статистически. Отсюда ясно, что статистическим расчётам надёжности нельзя придавать особенно большого значения, потому что далеко не все аварии имеет смысл изучать статистическими методами или говорить об их вероятности. Этот вывод важен для принципиальной оценки настоящего места теории надёжности, но сейчас наш интерес сосредоточен на другом, на выяснении, является ли статистически однородной наша тщательно отобранная совокупность.

Предположим, что кривая $p_2(t)$ даёт практически точное выражение вероятности аварии $p(t)$. Если аварии изоляции в основном вызываются её местными дефектами, то логично предположить, что авария данной машины не влияет, или мало влияет на её аварии после ремонта при дальнейшей эксплуатации. Но тогда общее число аварий на машине за всё время её эксплуатации есть сумма независимых случайных величин, числа аварий за время работы до исчерпывания времени эксплуатации машины. Каждое слагаемое имеет распределение Пуассона, следовательно, общее число аварий ξ_i на данной машине также имеет распределение Пуассона. Параметр числа аварий на машине i за время работы в $(k-1)$-й единице времени равен

$$\lambda_{ik} = p(t_k)S_i \approx p_2(t_k)S_i, \qquad (2.3.1)$$

где, как и раньше, S_i есть площадь изоляции i-й машины.

Поэтому параметр

$$\lambda_i = \mathrm{E}\xi_i \tag{2.3.2}$$

общего числа аварий i-й машины может быть вычислен суммированием выражений (2.3.1) по тем t_k, которые меньше срока эксплуатации данной машины. Последнее слагаемое, отвечающее неполному интервалу в 10^4 часов, определяется интерполяцией. И поэтому можно считать, что для всех машин нам известны числа (2.3.2).

Конечно, предложенный способ вычисления λ_i справедлив в предположении статистической однородности. Если она нарушается, то кривая $p_2(t)$ даёт лишь среднюю характеристику аварийности. Одни машины будут более аварийными, чем совокупность в среднем, другие менее аварийными.

Следовательно, на одних машинах будет больше аварий, чем полагается по закону Пуассона с параметром, вычисленным по нашему правилу, на других, меньше. Казалось бы, что мы и определили эффект, который следует искать для проверки нарушения статистической однородности.

Но беда в том, что этот эффект очень трудно заметить. Пусть, в самом деле, мы определили, что для некоторой машины $\lambda_i = 0{,}1$ в то время, как $\xi_i = 2$. Поскольку

$$P\{\xi_i \geq 2\} = \frac{\lambda_i^2}{2} e^{-\lambda_i} + \ldots \approx \frac{1}{200},$$

то, казалось бы, мы обнаружили значимое отклонение от статистической однородности. Но ничего подобного на самом деле нет. Ввиду того, что статистикой охвачено несколько сотен машин, для одной машины, и даже для нескольких, вполне может произойти событие с вероятностью 1/200.

Для того, чтобы сформулировать полезный статистический критерий однородности, можно пойти различными путями. Один из них состоит в том, чтобы вновь применить теорему Пуассона. Рассмотрим суммарное число машин с одной аварией, с двумя авариями, и т. д. Покажем, что мы можем вычислить распределение вероятностей для этих величин. Действительно, введём случайную величину

$f_k(\xi) = 1$, если $\xi_i = k$; и 0 в противном случае, $k = 1, 2, 3, \ldots$

Поскольку ξ_i есть число аварий на i-й машине, то число машин с k авариями равно $\sum f_k(\xi_i)$, сумме независимых случайных величин. Для большей части машин λ_i близко к нулю и поэтому при $k \neq 0$ вероятность

$P\{f_k(\xi_i) = 1\}$

мала. Это означает, что для указанной суммы приближённо действует распределение Пуассона. Его параметр подсчитывается таким образом:

$$\mathrm{E}[\sum_i f_k(\xi_i)] = \sum_i \mathrm{E} f_k(\xi_i),\ \mathrm{E} f_k(\xi_i) = P\{\xi_i = k\} = \frac{\lambda_i^k}{k!} e^{-\lambda_i}\ .$$

Здесь, если верна гипотеза статистической однородности, λ_i подсчитывается как сказано выше. Небольшое вычисление [1] показывает, что при разных k изучаемые суммы близки независимым случайным суммам.

В чём проявляется отклонение от статистической однородности? Одни машины будут более аварийными, с ними произойдут две или более аварии, другие, менее аварийными, с ними аварии вовсе не произойдут. Возрастёт число машин с двумя и более авариями, сократится число машин с одной аварией. Проявляется ли этот эффект в фактических данных? Вычисления приводят к следующим результатам: число машин с 1, 2, 3 и 4 авариями (ряд 1) и соответствующие ожидания (ряд 2)

1. 27 10 1 1
2. 29,6 5,7 1,5 0,44

Уменьшение числа машин с одной аварией по сравнению с его ожиданием незначительно, а увеличение их числа с двумя авариями заметно: для пуассоновского закона с параметром 5,7 вероятность значения 10 или больше равна 0,065. Для k = 3 и 4 отклонения не значимы.

Единственное отклонение, с которым стОит говорить, имеет место для машин с двумя авариями. Впрочем, можно смотреть на него как на самое сильное из четырёх независимых отклонений, и тогда его вероятность равна $1 - (1 - 0.065)^4 \approx 0{,}25$, и не особенно значимо.

Но в общем некоторая тень сомнения на статистическую значимость всё-таки брошена, хоть эта гипотеза почти что с честью прошла довольно жёсткое испытание. Это по-видимому означает, что для большинства машин аварийность примерно одинакова, хотя могут существовать небольшие выделяющиеся группы. Если бы аварийность разных генераторов была широко разбросана, то рассмотренный критерий дал бы гораздо более значимый результат.

Были исследованы различные конструктивные группы генераторов напряжением 10,5 квт. Оно проводилось близким методом, но значимость вычислялась иначе. Из 17 конструктивных групп были выделены три высокоаварийные (примерно 11% по числу машин) и не было выделено низкоаварийных групп. Таким образом, гипотеза статистической однородности полностью не подтверждается, но является вполне разумным приближением.

Следовательно, средней кривой $p_2(t)$ можно, в общем, пользоваться для приближённого вычисления среднего числа аварий для различных групп машин. Это обеспечивает критерий для оценки надёжности изоляции. Конечно, чисто статистические методы не имеют отношения к повышению надёжности работы, т. е. к технической задаче. Но по крайней мере мы можем сказать, что надёжность изоляции повысились или понизилась или осталась прежней. В этом практическое значение описанной работы. Она не имела бы такого значения, если не была бы установлена сравнительно высокая статистическая однородность изоляции. И тогда средняя кривая $p_2(t)$ дала бы лишь среднюю характеристику аварийности, от которой могли бы существовать любые отклонения.

2.4. Сравнение с глазомерным методом. Мы высказали предположение, что процедура сглаживания многочленом обычно успешна, потому что материал для такой обработки предварительно отбирается глазомерной оценкой. Было бы неэтично умолчать о том, что физики и инженеры часто выполняют на глаз и само сглаживание, не используя МНКв.

Но почему мы вообще решаем, что в том или ином случае сглаживание оказалось успешным? Не потому ли, что кривая, проведённая по МНКв, проходит как раз там, где мы провели бы кривую на глаз?

Был проделан эксперимент по глазомерному сглаживанию ломаной. В нём участвовали математики-сотрудники статистической лаборатории МГУ и инженеры – сотрудники ВНИИЭ. Каждому давался листок бумаги, на котором была изображена ломаная, но ничего больше, и предлагалось сгладить эту линию на глаз. После этого желающие могли наложить на свой чертёж кальку с кривыми $p_1(t)$, $p_2(t)$ и $p_3(t)$ и таким образом проверить свой глазомер, но вносить какие-либо изменения в проведённую кривую не разрешалось.

Глазомер подавляющего большинства участников оказался очень хорошим. Из 16 участников 15 почти полностью уложились в область между кривыми $p_1(t)$ и $p_3(t)$. Самый лучший результат этого опыта принадлежал И. В. Гирсанову, заведующему одним из отделов статистической лаборатории, который, к несчастью, погиб впоследствии в горном туристическом походе.

Один из участников эксперимента не уложился в указанную область только на начальном участке. В общем, по точности результаты глазомерного сглаживания вполне сравнимы с МНКв. Если бы нас интересовала только кривая $p_2(t)$, мы могли бы без всяких вычислений провести её на глаз. Но при аккуратной статистической обработке интересуются не только результатом, но и его точностью, а для этого уже нужна статистическая модель и вообще наука.

Например, при оценке вероятности успеха в испытаниях Бернулли мы в качестве оценки выбираем частоту успеха. Но, чтобы сообразить, как велико может быть отклонение частоты от вероятности, надо, во-первых, считать испытания независимыми (статистическая модель!) и, во-вторых, применить сложно доказываемую (при любом способе доказательства) теорему Муавра – Лапласа, т. е., бесспорно, науку.

При глазомерном сглаживании ломаной нам даже не нужно знать, по какому числу наблюдений получены точки этой ломаной, всё равно эти числа будет невозможно как-то использовать. Поэтому невероятно, чтобы кто-нибудь обладал способностью указывать на глаз доверительную область для искомой кривой $p(t)$. Здесь нужна вся наука, связанная с МНКв.

Тем не менее, почти полное совпадение результатов глазомерного сглаживания (области между 16 кривыми) с областью между кривыми $p_1(t)$ и $p_3(t)$) требует какого-то объяснения. Заметим, однако, что при малых t на начальном участке первая область несколько Уже, чем вторая, которая, как показывает расчёт, на начальном участке ещё примерно в 1,5 – 2 раза Уже, чем аккуратно построенная доверительная область с обычным коэффициентом доверия в пределах 0,70 – 0,95.

Итак, неопределённость глазомерного сглаживания по данному чертежу, в общем, всё же меньше неопределённости, которую полагается вычислять по правилам статистики. Дело в том, что при глазомерном проведении кривой мы имеем данный чертёж, т. е. определившийся результат случайных экспериментов, работая же статистическими методами, мы охватываем возможный разброс результатов случайных экспериментов при их различных реализациях.

В общем, однако, вопрос об истинных возможностях глазомерных методов требует широкого экспериментального исследования.

Глава 3. Теория случайных процессов

Как ни полезен в подходящих случаях МНКв, часто один взгляд на ряд наблюдений убеждает нас, что они не описываются моделью тренда с ошибкой. Мы не можем выделить на глаз детерминированную кривую, вокруг которой хаотически разбросаны наблюдения. Посмотрим, как писал об этом Слуцкий, один из основателей теории случайных процессов [8, с. 99]:

Почти все явления хозяйственной жизни, подобно многим другим процессам, социальным, метеорологическим и пр., протекают во времени чередой подъёмов и падений, подобно волнам, бегущим одна за другой. И, как на море последовательные волны не повторяют в точности друг друга, так и здесь соседние циклы никогда не совпадают одни с другими ни по продолжительности, ни по высоте подъёма.

И, однако же, как там, так и здесь, почти всегда сквозь всё это многообразие индивидуальных особенностей более или менее явственно проступают черты известного единообразия и некоторой приблизительной правильности. Глаз наблюдателя совершенно инстинктивно открывает на волнах одного порядка, другие, меньшие волны, так что идея гармонического анализа … кажется напрашивается сама собой.

Правда, *идею гармонического анализа* можно всё-таки попытаться осуществить с помощью модели тренда с ошибкой. Это делается так называемым *методом периодограммы*, который исторически предшествовал методам теории случайных процессов. От него в современную теорию перешло понятие *периодограммы*. Кратко рассмотрим этот метод.

3.1. Метод периодограммы. Предположим, что наши наблюдения в дискретные моменты времени, например, каждую секунду, могут быть описаны моделью

$$x_t = \sin(\lambda_0 t + \varphi) + \delta_t,\ t = 0, 1, ..., n, \tag{3.1.1}$$

где λ_0 есть некоторый параметр (круговая частота колебаний), φ, фаза колебаний, и δ_i, случайная ошибка. Предположим, что λ_0 много меньше 2π, так что последовательные наблюдения одной лишь составляющей $\sin(\lambda_0 t + \varphi)$ через единицу времени, которая много меньше периода колебаний $2\pi/\lambda_0$, давали бы отчётливо видную глазу синусоиду.

Прибавление случайных ошибок, пусть, для простоты, независимых друг от друга, конечно, испортит эту синусоиду. Спрашивается, как же восстановить частоту λ_0? Ответ заключается в следующем. Умножим наши наблюдения x_i на $\sin(\lambda t)$ и $\cos(\lambda t)$, где λ – переменное число (?), и образуем суммы

$$A(\lambda) = \sum_{t=1}^{n} x_t \sin \lambda t,\ B(\lambda) = \sum_{t=1}^{n} x_t \cos\lambda t.$$

Вычисление этих функций при разных и довольно часто выбираемых значениях λ было довольно громоздкой задачей, но с появлением ЭВМ эти вычисления перестали представлять трудности.

Образуем функцию

$$C(\lambda) = A^2(\lambda) + B^2(\lambda),$$

которая называется периодограммой. Раньше периодограмму представляли себе не как функцию частоты, а как функцию периода 2π/λ, что, конечно, безразлично, но объясняет, откуда взялся термин *периодограмма*.

Основное утверждение состоит в том, что при достаточно большом числе наблюдений n, периодограмма как функция λ будет иметь отчётливо выраженный максимум в малой окрестности истинной частоты λ_0.

Если детерминированная часть наших колебаний состоит не из одной гармоники $\sin(\lambda_0 t + \varphi)$, а из нескольких, т. е.

$$x_t = \sum_{j=0}^{k} A_j \sin(\lambda_j t + \varphi_j) + \delta_t, \tag{3.1.2}$$

то на периодограмме будет несколько максимумов, близких к точкам $\lambda_0, \ldots, \lambda_k$. Их высота будет зависеть от числа наблюдений n и амплитуд A_j. Здесь мы попадаем в довольно трудное положение, если заранее не знаем числа гармоник k и дисперсии σ^2 ошибок δ_i. На периодограмме, вообще говоря, очень много локальных максимумов, и непонятно, как их истолковывать, то ли как действительно отвечающие *скрытым* периодам λ_j, то ли как появившиеся чисто случайно.

Но с этими трудностями можно как-то справляться. Хуже другое. Как правило, ничто не гарантирует нам, что верна модель (3.1.2). Мы скорее можем быть уверены в обратном. Например, при исследовании рядов наблюдений, взятых из экономики, на глаз видно, что наблюдения довольно гладко зависят от времени (редко меняют возрастание на убывание и наоборот), чего не должно быть для наблюдений, выражаемых этой моделью, потому что они должны быть хаотически разбросаны вокруг плавной кривой

$$\sum_{j=0}^{k} A_j \sin(\lambda_j t + \varphi_j).$$

Можно, конечно, допустить, что эта кривая сама мало отвечает нашему представлению о плавной кривой и что её *негладкость* компенсирует случайный разброс, но для этого потребуется неразумно большое число гармоник.

Поэтому вопрос первостепенной важности состоит в том, насколько разумные результаты даёт метод периодограммы в том случае, когда модель (3.1.2) неверна, а ряд наблюдений $\{x_i\}$ описывается какой-то другой моделью. Такие экспериментальные исследования предпринял известный английский статистик М. Г. Кендалл. Они опубликованы в его довольно редкой книге [13].

Эта небольшая книжка является одной из самых замечательных книг по математической статистике. Любопытен эпиграф, с которого она начинается. Его можно перевести примерно так:

Основы наших знаний имеют мало общего с солидным фундаментом обычного здания. Их нужно скорее сравнивать со сваями домов Роттердама, которые как-то покоятся на глубоком ложе мягкой грязи.

В начале этой книги мы говорили о том, что *науки о благовониях* возникают, расцветают, но долго не держатся. Так и случилось с методом периодограммы. Погубил его, в частности, Кендалл своей книгой [13].

Он рассматривает модель авторегрессии (мы скоро будем говорить о ней), которая столь же применима к анализу экономических рядов, как и модель (3.1.2), если не более. И вот, для модели авторегрессии метод периодограммы выделяет частоты, совершенно не имеющие отношения к структуре модели. Заключение Кендалла об этом методе формулирует кратко: *As misleading as it could be* (приводит к настолько неверным результатам, насколько это вообще возможно).

В частности, по всей видимости именно таково отношение Кендалла к работам известного английского экономиста У. Бивериджа. Он знаменит тем, что составил и проанализировал методом периодограмм несколько длинных экономических рядов. Например, он составил и исследовал ряд цен на пшеницу в Европе, охватывающий 370 лет (хотелось бы знать, из каких соображений и верно ли составлен этот ряд, но этого, по-видимому, нельзя выяснить[3.1]). Составив периодограмму, Беверидж выделил много периодов в этом ряду, которые скорее всего не имеют смысла.

3.2. Случайные процессы. В настоящее время вместо метода периодограммы употребляется корреляционная и спектральная теория стационарных случайных процессов. Случайные процессы это функции переменного t, часто (но не обязательно) играющего роль времени и элементарного события ω. Мы будем сокращённо обозначать случайный процесс через ξ_i, опуская случайный аргумент ω. Дело в том, что функциональная зависимость случайного процесса от ω никогда не рассматривается в приложениях.

Если бы мы пожелали явно описать пространство элементарных событий Ω = {ω}, то, как правило, отдельные элементарные события были бы чрезвычайно сложными. Например, под таким событием часто понимают функцию $\omega = \omega(t)$ аргумента t. В этом случае значение случайного процесса в момент t на элементарном событии ω есть $\omega(t)$, что является чистой тавтологией и практически никуда не ведёт. Хотя указанное понимание случайного процесса необходимо для развития аксиоматической теории, практически использовать его не приходится. В приложениях всегда говорят лишь о распределениях вероятностей процесса ξ_i в какие-то моменты t_1, t_2, ..., t_n. Возможны два случая, когда время t принимает дискретные значения (наблюдения производятся в дискретные моменты времени) и когда t принимает непрерывные значения из некоторого промежутка.

Понятие процесса с непрерывным временем требует очень осторожного обращения[3.2]. Если принимать соответствующие математические теоремы слишком всерьёз, реализации таких процессов часто станут обладать парадоксальными свойствами, которые могут направить мысль исследователя по ложному пути. В брошюре [9] говорится о парадоксе, связанным с тем, что математическая модель броуновского движения обладает свойством, позволяющим точно определять коэффициент диффузии по наблюдениям сколь угодно малого участка реализации броуновского движения. (Этот пример см. в [9, § 2.5].)

Рассмотрим ещё один пример такого рода. Как известно, любая радиостанция работает в полосе частот ограниченной ширины. Если радиосигнал рассматривать как случайный процесс, то его спектр будет заключён в этом конечном интервале. И вот, существует математическая теорема, согласно которой с вероятностью 1 реализация ξ_i такого процесса является аналитической функцией *t*. Но такие функции полностью определяются своими значениями на любом сколь угодно малом интервале времени. Следовательно, послушав радиостанцию сколько угодно короткое время, мы можем однозначно определить, что она передавала в прошлом и будет передавать в будущем. Вывод откровенно абсурдный.

Конечно, не так трудно указать здесь ошибку. Во-первых, радиопередача не есть случайный процесс, это не элемент статистического ансамбля. Во-вторых, аналитическую функцию мы можем восстановить по её значениям на любом участке только в том случае, если знаем её там совершенно точно, тогда как функцию непрерывного аргумента нельзя знать точно во всех точках (точно нельзя записать даже одно число, тем более совокупность бесконечного множества чисел). В третьих, радиосигнал не есть аналитическая функция времени, потому что в XIX веке не было радиостанций, все радиосигналы равнялись нулю, а аналитическая функция, равная нулю на некотором отрезке, равна нулю всюду.

Здесь уместно отступление о понятии функции в математике[3.3]. При зарождении математического анализа под функцией обычно понималась формула, задающая зависимость $y = y(x)$. Конечно, за исключением немногих точек все функции были непрерывными и дифференцируемыми, так что вопрос о доказательстве взять производную [дифференцируемости] даже и не возникал, Но позже, уже в XIX веке, установился взгляд на функцию просто как на соответствие между множествами значений аргумента *x* и функции $y = y(x)$. Обычно требуется, чтобы каждому значению *x* отвечало ровно одно значение *y*, но, вообще говоря, не обратно. Конечно, произвольному соответствию $y = y(x)$ нет причины быть непрерывным или дифференцируемым.

Довольно трудно, а потому и интересно построить пример непрерывной, но нигде не дифференцируемой функции. Первый такой пример принадлежит Вейерштрассу, позже были придуманы другие, более простые. Объекты такого рода оказались очень интересными для математиков, и на них в значительной степени переключилось внимание.

Для нас особенно интересно то обстоятельство, что математические соображения теории случайных процессов приводят к тому, что реализации многих из них нужно считать непрерывными, но не дифференцируемыми функциями, или функциями, допустим, дифференцируемыми два раза, но такими, что вторая производная непрерывна, но уже не дифференцируема. Это было большой радостью, потому что вроде бы доказывало, что не дифференцируемые функции в самом деле существуют в природе.

Но мы хотим бросить тень сомнения на эту радость. Совершенно абсурдно считать, что не дифференцируемая функция может наблюдаться в эксперименте. Ведь такую реализацию случайного процесса нельзя задать ни формулой, ни таблицей, ни графиком, ни алгоритмом вычисления. Естественно, что если мы будем считать её действительно реальной, т. е. точно известной во всех точках, то сможем получить абсурдные выводы. То же относится к реализации таких процессов, которым полагается быть аналитическими функциями.

Мы будем рассматривать случайные процессы с дискретным временем, которые не таят в себе таких парадоксов, как процессы с непрерывным временем. В общем, для процесса с дискретным временем обычно можно считать, что наблюдаемые значения ξ_i известны точно.

Сделаем одно терминологическое замечание. В русской литературе существует термин *временной ряд*. Обычно он обозначает случайный процесс, часто – стационарный случайный процесс (определение см. в гл. 1). Но в английской литературе соответствующий термин *time series* обозначает значения любой, том числе неслучайной переменной, зависящей от времени и наблюдаемой в дискретные моменты времени. Такой объект мы будем называть *рядом наблюдений*. Во избежание смешения случайных и неслучайных величин мы не будем употреблять термина *временной ряд*, говоря вместо него либо *случайный процесс*, если предполагается случайность, либо *ряд наблюдений*, если случайность может быть, а может и не быть.

По причинам, изложенным в гл. 1, основную роль играют стационарные случайные процессы. Концепция случайного процесса позволяет думать о совместном распределении случайных величин ξ_i, но обычно ограничиваются двумерными распределениями, причём изучают только ожидание и корреляционную функцию

$$m(t) = \mathrm{E}\xi_i,\ B(s, t) = \mathrm{E}[(\xi_s - m(s)][(\xi_t - m(t)].$$

В случае стационарного процесса распределения вероятностей не меняются при сдвиге времени. Следовательно,

$$m(t) = \mathrm{E}\xi_i = m \qquad (3.2.1)$$

не зависит от t, а корреляционная функция зависит лишь от разности аргументов $(t - s)$:

$$B(s, t) = B(t - s). \qquad (3.2.2)$$

Процесс, удовлетворяющий лишь этим двум условиям, называется *стационарным в широком смысле*. Это и есть основное понятие, с которым современная математическая статистика до последнего времени рекомендует подходить к рядам наблюдений.

Теория случайных процессов, которая имеет дело только со средним значением и корреляционной функцией, называется *корреляционной теорией*. Мы сейчас рассмотрим её.

3.3. Корреляционная и спектральная теории. Основным достижением общей теории стационарных случайных процессов является теорема, устанавливающая, что в общем случае корреляционная функция представима в виде

$$B(s,t) = B(t-s) = \int_{-\pi}^{\pi} \cos\lambda(t-s)dF(\lambda)\,, \qquad (3.3.1)$$

где $F(\lambda)$ ограниченная неубывающая функция. Обычно предполагается, что существует *спектральная плотность*, т. е. функция $f(x) \geq 0$, такая, что

$$dF(\lambda) = f(\lambda)d\lambda, \text{ и тогда } B(t-s) = \int_{-\pi}^{\pi} \cos\lambda(t-s)f(\lambda)d\lambda. \quad (3.3.2)$$

Поэтому спектральный анализ, т. е. экспериментальное определение спектральной плотности $f(\lambda)$, иногда объясняется как определение дисперсий отдельных случайных составляющих процесса.

Для практического использования корреляционной или спектральной теории нужны две вещи: во-первых, надо знать, какие практические выводы можно сделать из корреляционной функции или спектра (спектральной плотности); во-вторых, надо уметь оценивать корреляционную функцию или спектральную плотность по наблюдениям.

Корреляционная функция имеет стандартное применение в статистических вопросах. Например, дисперсия среднего арифметического $\overline{\xi}$ выражается через сумму попарных ковариаций, т. е. через корреляционную функцию или через спектральную плотность. Но иногда оценка спектра или корреляционной функции применяется в качестве магического средства, которое якобы даёт возможность проникнуть в природу наблюдаемого процесса. Надо отчётливо представлять себе, что корреляционная теория, вообще говоря, оперирует с такими характеристиками, которые далеко не определяют процесса в целом и часто дают лишь поверхностные сведения о нём.

Если нас интересует некий вопрос о структуре процесса, мы должны быть в состоянии сформулировать его в терминах корреляционной теории, учитывая притом, что ни корреляционной функции, ни спектральной плотности мы обычно не знаем точно, а оцениваем по наблюдениям. Таким образом, речь должна идти о сравнительно грубых характеристиках, которые можно определить по неточным данным.

Например, в литературе существует так называемый *метод канонических разложений*, для применения которого полагается найти собственные функции интегрального уравнения, в которое в качестве ряда входит корреляционная функция процесса. Этот метод следует признать практически безнадёжным, потому что неточность ядра интегрального уравнения весьма существенно отражается на собственных функциях. Действительно, нам неизвестно каких-либо практических применений метода канонических разложений. Все так называемые *применения* исходят из произвольного задания корреляционной функции и не имеют дела со статистическим материалом.

Оценка корреляционной функции и спектра довольно сложна. Сначала полагается оценить и вычесть среднее значение процесса. Оценкой для *m* является среднее арифметическое: $\hat{m} = \overline{\P}$, а оценка

$$B(u) = \mathrm{E}\xi_t\,\xi_{t+n} - m^2$$

равна

$$\hat{B}(u) = \frac{1}{n-u}\sum_{t=1}^{n-u}\xi_t\xi_{t+u} - \hat{m}^2,\ u = 0,\ 1,\ ...,\ n-1.$$

Она обладает рядом неприятных свойств. Во-первых, для эргодического процесса фактические значения $B(u)$ быстро убывают с ростом u. Но стандартные отклонения оценок $\hat{B}(u)$ примерно одинаковы для всех u, имея порядок $1/\sqrt{n-u}$. В результате при u порядка нескольких десятков сами величины $B(u)$ очень малы, составляя сотые и тысячные доли от $B(0)$, а стандартные отклонения при не слишком большом n составляют десятые доли от $B(0)$, так что оценка не имеет никакого смысла.

Во-вторых, эти оценки $\hat{B}(u)$ при близких значениях u не разбросаны хаотически около истинных значений, так как соседние оценки $\hat{B}(u)$, $\hat{B}(u+1)$, $\hat{B}(u+2)$, ... коррелированы друг с другом. При взгляде на график значений $\hat{B}(u)$ глаз автоматически выделяет довольно правильные колебания при неразумно больших значениях u, тогда как в действительности $B(u)$ неотличимо от нуля. Итак, при оценках корреляционной функции нельзя доверять глазомеру, и все наши действия оказываются неуверенными.

Оценка спектральной плотности $f(\lambda)$ предпочтительней. Если оценивать её в точках $\lambda = \lambda_1, \lambda_2, ..., \lambda_m$, не расположенных слишком близко друг к другу, то соответствующие оценки $\hat{f}(\lambda_i)$ окажутся почти независимыми случайными величинами. Этот факт впервые открыл Е. Е. Слуцкий.

Для оценки спектральной плотности применяется всё та же подходяще нормированная периодограмма. Но она очень изрезана, потому что её дисперсия не стремится к нулю с ростом числа наблюдений. Поэтому применяется сглаживание с тем или иным весом, т. е. усреднение периодограммы. В результате получается оценка не самой спектральной плотности, а той функции, которая получится, если эту плотность усреднить с тем же весом. Значит, интервал усреднения надо выбирать небольшим. Но усреднение по небольшому интервалу мало уменьшит дисперсию периодограммы и практические рекомендации являются результатом компромисса между этими противоречивыми требованиями к усреднению.

Мы не можем входить в детали математических приёмов. Оценка корреляционной функции и спектра в учебниках по теории случайных процессов обычно не излагаются сколько-нибудь удовлетворительно. Учебники, как мы говорили, предпочитают исходить из предположения, что случайный процесс задан вместе со своей корреляционной функцией. Весьма надёжным источником сведений по статистическим вопросам можно назвать, например, переводную книгу [10], но она написана весьма сжато и читается с трудом. Недавно переведена книга [3]. Читать её легче, но она менее надёжна. Так, в ней не сказано достаточно отчётливо, что все изложенные для дисперсий оценок корреляционной функции и спектральной плотности формулы применимы вовсе не к любому стационарному процессу (требуются сильные условия, выражающие в математической форме свойство эргодичности). Но книгой всё же можно пользоваться, хотя предлагаемые *практические* примеры требуют весьма критического отношения.

Мы хотим предупредить читателя, который обратится к указанной литературе, о следующем. Материал, на котором складывались методы теории случайных процессов, в основном брался из экономических данных и обычно очень невелик. Изменение некоторого экономического показателя можно проследить за десятки, в лучшем случае за несколько сотен лет, как в случае ряда Бивериджа. Один год обычно обозначает одно наблюдение, иначе вмешается сезонная периодичность, с которой придётся как-то разбираться, и вообще большинство экономических показателей вычисляется за целый год.

Итак, мы имеем десятки или сотни наблюдений, а расчёты показывают, что для надёжной оценки корреляционной функции или спектральной плотности необходимы тысячи и десятки тысяч наблюдений. Этот вывод сделал уже Кендалл [13] в отношении корреляционной функции. В результате математики стараются чего-то добиться выбором оптимального способа сглаживания периодограммы, но если наблюдений мало, то это, в общем, безнадёжно.

Настоящая сфера приложения теории случайных процессов там, где можно иметь сколько угодно наблюдений. Радиофизики давно разработали методы, которые позволяют легко и просто получать оценку спектра случайного процесса, если только не нужно экономить на количестве наблюдений. Они применяют системы фильтров, выделяющих отдельные полосы частот [6, часть 2, с. 14 – 15].

3.4. Обзор практических применений. Среди основателей теории случайных процессов, работавших также со статистическим материалом, должны быть названы Э. Юл, Е. Е. Слуцкий и М. Г. Кендалл. Важнейшее значение имеют работы Колмогорова, о которых мы будем говорить позже. Эти статистики работали ещё до второй мировой войны без применения каких-либо автоматизированных средств обработки, и, естественно, могли иметь дело, самое большее, с сотнями наблюдений. Тысячи и десятки тысяч требуются в корреляционной теории потому, что мы хотим слишком много определить из них: не оценку одного или нескольких параметров, а бесконечно много величин $B(u)$, $u = 0, 1, 2, ...$, т. е. корреляционную функцию, или спектральную плотность, которая является функцией $f(\lambda)$ аргумента λ, пробегающего отрезок $[-\pi, \pi]$.

Чтобы сделать методы корреляционной теории практически эффективными при небольшом числе наблюдений, можно пойти по другому пути, а именно искать модели, зависящие от малого числа параметров. Одну такую модель предложил Слуцкий, модель скользящего среднего.

Представим, что имеется бесконечная последовательность независимых одинаково распределённых случайных величин

$$... \xi_{-1}, \xi_0, \xi_1, ..., \xi_n, ...,$$

но что мы наблюдаем не её, а последовательность

$$... \varsigma_{-1}, \varsigma_0, \varsigma_1, .., \varsigma_n, ..., \quad (3.4.1)$$

где

$$\varsigma_n = \sum_{k=0}^{m} \alpha_k \xi_{n-k}. \quad (3.4.2)$$

Иными словами, ς_n это сумма некоторого числа независимых величин ξ_{nJk}, умноженных на подходящие α_k. В опытах Слуцкого моделировалась последовательность $\{\xi_n\}$, а для получения ς_n на строчку модулированных чисел накладывалась рамка с окошечками, в которые были видны ξ_n, ξ_{nJ1}, ..., ξ_{nJk}. Для получения ξ_{n+1} рамка передвигалась на одну позицию вправо, отсюда и название, *скользящее среднее*. Параметрами являлись числа α_0, ..., α_m.

Слуцкий показал, что модель скользящего среднего может дать картину волнообразных колебаний, очень похожих на колебания экономических показателей. Но такого утверждения, что все статистические свойства какого-либо ряда наблюдений, взятого из практики, описываются моделью скользящего среднего, у Слуцкого нет. Насколько нам известно, в аккуратных статистических работах такого рода примеров вообще нет[3.4].

Здесь необходимо сказать, что статистическая работа с рядами наблюдений требует разносторонних статистических проверок принятой модели. Это понимал Слуцкий, понимали и представители серьёзной английской школы как Юл и Кендалл. Но, к сожалению, такое понимание в настоящее время утрачено, если, конечно, иметь в виду среднюю публикацию по приложениям случайных процессов. Для общей массы публикаций характерна убеждённость в том, что теория случайных процессов обязана работать во всех случаях жизни. В результате основные предпосылки, главная из которых – статистический, а не просто неопределённый характер самогО явления, никак не проверяются. Появляется поток публикаций, которые просто не заслуживают серьёзного рассмотрения. Для серьёзного анализа остаётся совсем немного работ.

Из числа этих последних в первую очередь назовём работу Юла [17]. Она исследует изменение во времени числа солнечных пятен. Юл прежде всего отвергает модель периодограммы, потому что для периодограмм случайность присуща лишь ошибкам наших измерений и совершенно не влияет на ход самого процесса.

Юл замечает, что следует иметь какую-то модель процесса, в котором случайное вмешательство в данный момент времени влияет на дальнейшее проведение процесса. Представим, например, что мы наблюдаем качание маятника, но какие-то мальчишки принялись стрелять в него горошинами[3.5]. Каждый случайный удар горошины изменяет скорость маятника, а потому влияет на весь последующий процесс. Правда, от мальчишек трудно ожидать статистически однородной стрельбы горохом, но в реальных процессах, таких, как солнечная активность или экономическая жизнь, возможно, что иногда существует статистическая однородность случайного вмешательства.

Будем наблюдать положение маятника в дискретные моменты времени, но достаточно часто, чтобы на одном периоде невозмущённых колебаний было много наблюдений. Получим последовательность наблюдений

$$\xi_0, \xi_0, ..., \xi_n, ...$$

и Юл полагает, что она может быть описана моделью вида

$$\xi_n + a\xi_{n-1} + b\xi_{n-2} = \delta_n. \qquad (3.4.3)$$

Здесь a и b это постоянные числа (?) (параметры модели),$\{\delta_n\}$, последовательность одинаково распределённых независимых величин, таких, что δ_n не зависит от ξ_{n-1}, ξ_{n-2}, ... Считается, что $E\delta_n = 0$, а $\sigma^2 = D\delta_n$ является третьим параметром модели.

Это – знаменитая модель авторегрессии (второго порядка), с которой работали многие заслуживающие полного доверия статистики.

Правда, в случае маятника рассуждения Юла, приводящие к модели (3.4.3), не совсем ясны. Анализ показывает, что они и неверны. В частности в случае маятника $\{\delta_n\}$ не есть последовательность независимых случайных величин, она описывается скользящим средним по Слуцкому. Но внести в модель ещё и параметры скользящего среднего означает иметь слишком много параметров и крайне усложнить работу с моделью.

Ошибка Юла в случае маятника, разумеется, не доказывает логически, что модель авторегрессии (3.4.3) неприменима к солнечным пятнам или к каким-то экономическим показателям, но, конечно, это дурное предзнаменование.

Как писал Декарт, мир можно объяснить многими способами, но вопрос лишь в том, чтобы выбрать тот самый, который на самом деле правилен. Наибольшие шансы оказаться правильным имеет, разумеется, самый естественный и гармоничный и, в частности, не содержащий противоречий. Если же при самом зарождении теории её создатель допустил ошибку, хотя бы в каком-то частном случае, то наши надежды на успех в других случаях резко уменьшаются.

Что касается солнечных пятен, то сам Юл не получил решительного положительного результата. Он был вынужден внести изменения в модель (3.4.3), предположив, что мы наблюдаем не сами величины $\{\xi_n\}$, описываемые этой моделью, а что наши наблюдения испорчены дополнительной случайной ошибкой. Это пришлось допустить потому, что, как и полагается, Юл подверг свою модель статистической проверке, которую она не выдержала. Изменение модели позволяет свести концы с концами, но за счёт введения дополнительного параметра, а это в статистике очень плохо[3.6]. В общем, [17] – пример образцовой статистической работы, давшей, однако, сомнительный, если не отрицательный результат, что часто бывает именно с образцовыми работами.

В связи с возникновением интереса к прогнозу случайных процессов другой представитель английской школы, П. А. П. Моран [15], исследовал возможности модели (3.4.3) для прогноза солнечной активности. Поскольку δ_n не зависит от прошлого процесса, т. е. от величин ξ_{n-1}, ξ_{n-2}, ..., лучшее, что можно сделать для прогноза оценки величины ξ_n по всему прошлому, состоит в том, чтобы положить, что

$$\hat{\xi}_n = -\,(a\xi_{n-1} + b\xi_{n-2}).$$

Моран так и сделал и показал результаты своим друзьям-радиофизикам. Они сказали, что прогноз такого качества можно сделать без всякой науки, просто на глаз. Поставили эксперимент, и так оно и оказалось. Такова была вторая неудача модели авторегрессии.

Но эта модель обладает тем прекрасным качеством, что с ней очень легко работать. Параметры модели легко оцениваются, корреляционная функция имеет вид затухающих синусоидальных колебаний, так что её сравнительно легко истолковывать. Простой вид имеет и спектральная плотность. О работе с моделью авторегрессии лучше всего прочесть в книге [14]. Поэтому имело смысл много раз испробовать эту модель на различных материалах в надежде найти случаи, когда она будет работать хорошо.

Ещё до Морана моделью авторегрессии много занимался Кендалл. Он ограничивался такими значениями параметров *a* и *b* в формуле (3.4.3), при которых она определяет стационарный процесс и притом в основном экономическими рядами. Такие ряды редко колеблются около одного какого-то уровня, образуя стационарный процесс. Обычно они имеют тенденцию, *тренд*. Например, производство электроэнергии возрастает, в общем, экспоненциально. Следовательно, в логарифмической шкале оно имеет линейный тренд. Вопрос в том, чтобы описать отклонения от общей тенденции, происходящие в разные годы.

Кендалл полагал возможным определить и вычесть тренд тем или иным способом, исключая, конечно, глазомерный как слишком субъективный для строгой статистической школы. Это вносит дополнительные усложнения в статистическую структуру остающихся отклонений от тренда, но делать нечего. Именно такие отклонения, вроде бы образующие стационарный процесс, и изучались с помощью модели авторегрессии.

Трудно высказаться определённо о результатах Кендалла. В одних случаях статистические проверки благополучно выходили, в других случаях – нет. Можно ли считать, что иногда был достигнут успех, или же следует объяснить благополучие статистических проверок малым числом наблюдений? Так или иначе, нет никакого объяснения, почему, скажем, освобождённый от тренда ряд цен на пшеницу не подходит под модель авторегрессии[3.7], а поголовье овец подходит. И сколько-нибудь решительного успеха в обработке экономических рядов не достигнуто.

Кендалл [13] занимался исследованием процессов авторегрессии, построенных по уравнению (3.4.3), с помощью таблиц случайных чисел. Самый длинный из моделированных рядов имеет 480 членов. Посмотрим в заключение на вид эмпирической оценки корреляционной функции по этим данным. Обратим внимание, как сильно оценка отличается от истинных значений, затухая гораздо медленнее истинной корреляционной функции (рис. 3).

Оценку спектральной плотности ряда Кендалла опубликовал Хеннан [10]. Графики теоретической плотности и её различных оценок изображены на рисунке 4. Видно, что оценки довольно мало похожи на истинную плотность. В частности, последняя имеет максимум близко к точке $\lambda = \pi/5$, тогда как максимумы оценок лежат в точке $\lambda = \pi/15$.

Непривычному глазу может показаться, что малые значения спектральной плотности оцениваются хорошо. Но на самом деле ничего подобного нет: относительная ошибка здесь так же велика, как и в левой части чертежа для больших значений спектральной плотности.

Мы видим, что корреляционная теория, созданная основателями теории случайных процессов для обработки дискретных рядов наблюдений, таких, как число солнечных пятен за разные годы или значения экономических показателей, как раз на этом материале не достигла бесспорных успехов. Идея математического описания волнообразных процессов столкнулась с тем практическим затруднением, что для сколько-нибудь приличных оценок корреляционной функции нужно иметь не десятки или сотни отдельных наблюдений, а, как указывает Кендалл [13], десятки и сотни волн, о которых идёт речь. Это превращается в тысячи и десятки тысяч наблюдений. С другой стороны, параметрические модели, в частности, модель авторегрессии, не нашли убедительного статистического подтверждения.

Следовательно, приложения теории случайных процессов на этом материале, в частности, прогнозирование, не имеют достаточного научного обоснования. Хуже всего то, что в этой области публикуется много работ (например, [4]), которые не применяют достаточных статистических проверок принятой модели, и потому не могут рассматриваться всерьёз.

Положение было бы совсем плохим, если бы одновременно с практической неудачей указанных работ не были созданы новые области приложений корреляционной теории в аэрогидродинамике и физике, которые и составляют её настоящую ценность. Мы и рассмотрим эти приложения.

3.5. Процессы со стационарными приращениями. Когда мы имеем какой-то математический аппарат и хотим подойти с ним к описанию явлений природы, то главное – не запрашивать у неё слишком много, не пытаться использовать этот аппарат в тех случаях, когда он бессилен. В частности, когда мы представляем себе некоторое волнообразно протекающее явление, нам хотелось бы применить для его описания теорию стационарных случайных процессов. Но постепенно был осознан тот факт, что самые крупные волны в наблюдаемом процессе могут либо вовсе не носить статистического характера, либо в наших наблюдениях может быть недостаточно данных, чтобы определить их статистические характеристики, либо, наконец, их чисто статистическое описание может быть недостаточно для наших целей.

Например, цикличность экономической жизни, видимо, обладает всеми этими тремя признаками. Мы здесь принципиально не можем рассматривать явление как статистическое, потому что нам дана одна реализация и нет никакого статистического ансамбля. И, конечно, наших наблюдений обычно недостаточно. Наконец, статистическое описание нас не удовлетворяет потому, что нам надо знать, например, не как развивается в среднем тот или иной спад или подъём, а как будет развиваться именно тот спад или подъём, который существует в данный момент. Таким образом, рассчитывать на описание экономических явлений самых крупных масштабов в рамках теории случайных процессов совершенно нельзя. Другое дело – явления малых масштабов, где, быть может, удастся что-то сделать.

Возьмём другой пример, протекание метеорологических процессов. Совершенно ясно, что статистическое описание самых крупных изменений погоды вековых масштабов невозможно и не имеет смысла. Априорно неясно, можно или нельзя применять статистические методы для описания изменения погоды в малых масштабах, т. е., например, в течение нескольких дней, с целью прогнозов погоды. Но практический опыт показывает, что это достаточно бесполезно. Если же ограничиться областью малых пространственных и временнЫх масштабов, то статистические методы дают блестящий успех. Соответствующая теория называется *статистической теорией турбулентности Колмогорова – Обухова*, и мы позже скажем о ней несколько слов.

Если теперь обратиться к геологии, и поставить, например, задачу об оценке запасов месторождения по данным о процентном содержании полезного компонента в ряде точек опробования, то и здесь обнаружится рискованность использования стационарных процессов для описания процентного содержания по всему месторождению в целом. Здесь опять плохо с ансамблем реализаций и недостатком данных для определения статистики самых крупных колебаний.

С другой стороны, самые крупные неоднородности имеют большой масштаб и вероятнее всего меняются плавно, и можно ожидать, что они неплохо отражаются нашей разведочной сетью, так что мы их знаем достаточно точно и в статистическом описании не нуждаемся. Но что же делать с неоднородностями малых масштабов, которые также могут влиять на оценку запасов?

Обратимся, наконец, к радиофизике – науке, где понятие случайного процесса пользуется наибольшим признанием. В этой науке всякого рода помехи и шумы считаются стационарными случайными процессами. Но среди шумов есть особый, фликкер-шум или эффект мерцания. Он объясняется хаотическими вариациями эмиссионной способности катодов электронных ламп. В литературе (см., например, [7]) достаточно чётко указывается, что эффект мерцания навряд ли можно описать моделью стационарного случайного процесса.

Итак, в настоящее время мы начинаем представлять себе, что математическое описание самых крупных волн волнообразно протекающих процессов методами математической статистики в большинстве случаев невозможно. Приходится рассчитывать на описание явлений меньших масштабов. Конечно, мы от многого при этом отказываемся.

Например, для прогноза погоды теория микроструктуры турбулентностей бесполезна, потому что она не описывает наиболее существенных крупномасштабных явлений. Но она полезна в других областях, например, при расчёте прохождения света через атмосферу, что важно для астрономии (учёт искажения изображения в телескопе).

Имеется одно универсальное понятие, введённое Колмогоровым понятие *процесса со стационарными приращениями*, которым во всех рассмотренных случаях можно надеяться заменить понятие стационарного случайного процесса.

В случае дискретного времени это означает, что от наблюдаемого процесса

$$\ldots\ \xi_{-1},\ \xi_0,\ \xi_1,\ \ldots,\ \xi_n,\ \ldots$$

мы переходим к разностям

$$\ldots\ \eta_{-1} = \xi_{-1} - \xi_{-2},\ \eta_0 = \xi_0 - \xi_{-1},\ \eta_1 = \xi_1 - \xi_0,\ \ldots,$$

которые и считаем реализацией стационарного случайного процесса. В случае непрерывного времени от наблюдаемого процесса $\xi(t)$ мы переходим к производной

$$\eta\,(t) = \xi'(t),$$

которую объявляем стационарным случайным процессом; иногда это дифференцирование надо понимать в обобщённом смысле[3.8].

Объясним подробнее, на что мы рассчитываем, переходя к разностям или производным. Представим себе, что наблюдаемый процесс есть сумма

$$\xi(t) = a(t) + \varsigma(t)$$

некоторой случайной или неслучайной составляющей $a(t)$, которая имеет характер крупных волн, и составляющей $\varsigma(t)$, которая меняется гораздо быстрее и которую рационально считать стационарным случайным процессом. Мы признаём, что описание изменений $a(t)$ не в наших силах и хотим изучать мелкомасштабные изменения, которые в основном определяет другая составляющая. Как раз это и достигается дифференцированием, потому что крупная составляющая $a(t)$ скорее всего меняется медленно, т. е. её производная мала. Имеем

$$\eta(t) = \xi'(t) = a'(t) + \varsigma'(t) \approx \varsigma'(t),$$

т. е. $\xi'(t)$ практически не содержит составляющей, связанной с $a(t)$. То же самое достигается разностью в случае дискретного времени.

Конечно, в случае непрерывного времени мы также можем не дифференцировать, а изучать разности

$$\Delta_\tau\xi(t) = \xi(t + \tau) - \xi(t) \approx \int_t^{t+\tau} \eta(s)ds\ .$$

Здесь $\eta(s)$ стационарный процесс. Последнее равенство служит для построения корреляционной и спектральной теории процессов со стационарными приращениями как интегралов от стационарного процесса.

Вместо корреляционной функции употребляется введённая Колмогоровым структурная функция

$$D(\tau) = \mathrm{E}[\Delta_\tau \xi(t)]^2,$$

т. е. дисперсия приращения процесса за время τ. С техникой работы с процессами со стационарными приращениями можно ознакомиться по книге [6, т. 2].

Таким образом, положение, сложившееся сейчас в науке, можно охарактеризовать так. Мы не ждём, что общие статистические методы могут описывать волнообразно протекающие процессы в целом, т. е. включая крупные волны. В общем, понятие стационарного случайного процесса скомпрометировано. Очередь за понятием случайного процесса со стационарными приращениями, которое не претендует на охват явления в целом, но может охватить его в области малых масштабов. Возможности этого понятия ещё не исследованы в достаточной мере. В затронутых нами примерах положение таково.

В области экономики существуют работы американской школы Бокса, см., например, [11] и [12]. Но статистическая доброкачественность подхода здесь сомнительна: статистических проверок не делается, внимание сконцентрировано на прогнозе, а между тем исключение крупномасштабной составляющей заставляет нас думать, что от прогноза вообще скорее следует отказаться. Ведь прогноз в первую очередь зависит именно от крупномасштабной составляющей. В общем, положение сомнительно, мы рассмотрим его позже.

Для метеорологии процессы со стационарными приращениями особого значения не имеют. Теория турбулентности Колмогорова – Обухова относится, собственно, к аэрогидродинамике. Здесь достигнут блестящий успех, подтверждены экспериментальные выводы, сделанные основателями теории. Впрочем, этот успех остаётся единственным.

В геологии имеется книга [5]. Её фактический материал в какой-то степени поддерживает ту гипотезу, что структурная функция содержания полезного компонента имеет вид

$$D(r) = \alpha \ln r + \beta,$$

где r это расстояние между точками опробования, α и β, параметры модели, определяемые из наблюдений. Книга, однако, содержит ряд несообразностей. Например, логарифмическая зависимость продолжается в область малых r, что невозможно, потому что $D(r)$ неотрицательная величина.

Далее, в одних случаях речь идёт о содержаниях, в других – о логарифмах содержания. Никаких статистических проверок не делается. И всё-таки фактический материал производит настолько сильное впечатление, что делает желательным аккуратные исследования в этом направлении, которым можно было бы доверять.

В радиофизике научный уровень высок, так что подобных несообразностей просто не может быть. Но здесь, насколько нам известно, нет никаких сообщений об успешном применении модели процесса со стационарными приращениями. В книге [7] высказана лишь гипотеза о том, что таким образом следовало бы описывать эффект мерцания.

В заключение остановимся подробнее на статистической теории турбулентности и на вопросах прогноза.

3.6. Статистическая теория турбулентности. Эта теория – блестящий успех чисто статистического описания явления. Речь идёт об описании картины очень развитой и очень сложной турбулентности с большим количеством вихревых движений разных масштабов. Основы теории создали перед войной работы Колмогорова и Обухова. Подтверждение сделанных теоретических выводов на опыте потребовало совершенной измерительной техники и заняло около 25 лет. Разрабатываются и применения этой теории к вопросам распространения электромагнитных и акустических колебаний в атмосфере.

Понятно, что точное знание поля скоростей в турбулентном потоке и невозможно, и бесполезно. Действительно, если бы мы и имели какой-то способ вычислять все скорости во всех точках, то одна их запись с достаточной точностью потребовала бы невообразимого количества бумаги или магнитной ленты. С таким количеством информации совершенно нельзя работать, речь должна идти о каком-то варианте статистического описания.

Оказалось, что основные понятия, пригодные для этой цели, можно взять из корреляционной теории. Но таких понятий в их первоначальном виде недостаточно. В науке существует свой закон: *из ничего не выйдет ничего*, т. е., если сам не придумаешь ничего нового, то за счёт применения уже сложившихся теорий ничего нового и не охватишь.

Постараемся, не вдаваясь в математические подробности, посмотреть, как конкретно действует этот закон в случае турбулентности, и какие новые соображения нужно было привлечь, чтобы сдвинуть дело с мёртвой точки. Представим себе турбулентный поток. Имеется средняя скорость течения, которая зависит от конкретных условий (что и где возбуждает поток – ветер в поле, течение в океане или компрессор в аэродинамической трубе). Эту скорость бессмысленно описывать статистическими методами. Но разности скоростей в различных точках потока и в разные моменты времени уже меньше зависят от граничных условий и в большей степени определяются свойствами самой жидкости или газа. Итак, изучаем разности

$$u(x_1, x_2, t_1, t_2) = v(x_1, t_1) – v(x_2, t_2),$$

где $v(x, t)$ скорость жидкости в точке x в момент t, причём точка x далека от границ потока, а время t достаточно велико, чтобы успел установиться стационарный режим.

Естественно предположить, что турбулентность стационарна в том смысле, что статистические характеристики разности u зависит лишь от разности $t_1– t_2 = τ$. Но остаются ещё трёхмерные переменные x_1, x_2, да и сама разность u есть разность скоростей, т. е. трёхмерный вектор. Итак, имеем трёхмерное поле векторов, зависящих от шести пространственных и двух временнЫх переменных, статистические свойства которого, впрочем, зависят лишь от разности между временнЫми переменными. Если на этом остановиться и ожидать от опыта определения статистических характеристик такого поля, то опыт с этим никогда не справится, ибо это практически невозможно. Наука зашла в тупик.

Именно таково положение в некоторых других науках. Можно вводить случайные тензоры напряжений, случайную прочность, упругость и т. д., но польза этих понятий равна нулю, так как нельзя определить их статистические характеристики. Очевидно, в теории турбулентности необходимо было дальнейшее теоретическое развитие. Прежде всего, в достаточно развитой турбулентности все точки и все направления должны быть равноправны. Это утверждение кажется простым, на самом же деле оно довольно тонко. Можно представить себе измерительный прибор, состоящий из трёх векторов (x, e_1, e_2), причём последние два приложены к началу вектора x и все три вектора жёстко соединены. Измерение состоит в том, что мы прикладываем вектор x его началом к точке x_1 потока. Его конец будет в точке потока $x_2 = x_1 + x$. Мы берём проекцию разности скоростей $v(x_2, t) – v(x_1, t)$ на направления e_1 и e_2 и получаем две случайные величины.

В корреляционной теории наблюдаемой считается их корреляция. Так вот, эта корреляция не должна меняться ни при каких вращениях тройки (x, e_1, e_2) как твёрдого тела и не должна зависеть от точки x_1. Турбулентность, удовлетворяющая этому условию, называется локально *изотропной*. Можно показать, что в несжимаемой жидкости при изотропной турбулентности все статистические характеристики векторного поля $u(x_1, x_2, t_1, t_2)$ выражаются через характеристики его любой составляющей, т. е. через проекции поля на любую ось координат. Можно считать, что x_1 и x_2 лежат на этой оси. Итак, задача свелась к одной случайной функции двух одномерных пространственных переменных и двух временнЫх.

Свести дело к каким-нибудь одним переменным, либо пространственным, либо временнЫм, помогает *гипотеза замороженности*. Она состоит в том, что турбулентные *завитушки* переносятся основным потоком без изменений, будто *вмороженными* в жидкость. В таком случае нам нет надобности измерять турбулентность в разных точках x_1, x_2. Мы расположим линию (x_1, x_2) вдоль скорости основного потока, поставим прибор в точку x_2 и подождём, пока турбулентность из точки x_1 не перенесётся в точку x_2. Всё свелось к функциям одного лишь времени.

Гипотеза замороженности проверялась экспериментально, хотя, строго говоря, не для самОй турбулентности, а для её статистических характеристик, и хорошо согласовалась с опытом.

Сведя дело к одной функции пространства или времени, т. е. к обычному процессу со стационарными приращениями, уже можно было на что-то рассчитывать. И всё же для опытного определения остаётся структурная функция, что ещё слишком много. Нужен параметрический вид этой функции $D(r)$, где r расстояние между точками, в которых измеряется составляющая скорости.

Здесь важнейшие соображения высказал Колмогоров. В соответствии с ними $D(r)$ может зависеть лишь от вязкости жидкости, которая ответственна за диссипацию, т. е. за превращение в тепло энергии турбулентных неоднородностей (и таким образом за *гашение* турбулентности) и от количества энергии, которое, заимствуясь у основного потока, постепенно передаётся от больших вихрей к малым (и тем самым *поддерживает* турбулентность, энергия же, конечно рассчитывается на единицу массы жидкости и времени). Следовательно,

$$D(r) = \varphi(r, v, \overline{\varepsilon}),$$

где φ некая универсальная функция, v и $\overline{\varepsilon}$, параметры, причём вязкость v известна, а количество энергии $\overline{\varepsilon}$ единственный параметр, изменяющийся от эксперимента к эксперименту.

Если расстояние r достаточно мало по сравнению с размерами потока, чтобы можно было пользоваться моделью изотропной турбулентности, но настолько велико, что для вихрей размера r вязкость ещё несущественна, то $D(r)$ не зависит от v. Из соображения подобия в этом случае оказывается, что

$$D(r) = C\overline{\varepsilon}^{2/3} r^{2/3}, \qquad (3.6.1)$$

где C универсальная постоянная.

Для меньших r, при которых вязкость существенна, формулы для $D(r)$ не найдено, однако известно, что

$$D(r) = (v\overline{\varepsilon})^{1/2} \beta[\frac{r}{(v^3\overline{\varepsilon})^{1/4}}],$$

где β некоторая универсальная функция одного переменного. Зависимость (3.6.1) называется *законом двух третей Колмогорова*. Существуют спектральные аналоги всех утверждений о структурной функции.

В 1940 – 1941 гг., когда были опубликованы указанные выводы, всё это было гипотезой. После войны началась их интенсивная экспериментальная проверка. Структурные функции очень похожи на корреляционные, так что их оценки имеют те же неприятные свойства. Удобнее было проверять путём эмпирических измерений спектров. Конечно, никто не вычислял сглаженной периодограммы, а применялись фильтры, подробности см. в [6]. Мы же будем кратки и скажем просто, что измерения подтвердили всё, и закон двух третей в достаточно широком интервале значений r, и универсальность постоянной C, и универсальную зависимость $D(r)$, выраженную через функцию β для малых значений r. Рассуждения, основанные на здравом смысле и теореме размерности, оказались исключительно удачными, хоть они и не могут быть совершенно точными, потому что против них имеются некоторые физические соображения. Вся теория имеет чисто статистический характер; её замысел – охватить довольно грубыми соображениями основные черты статистики изучаемого явления, чтобы стала возможной экспериментальная проверка. Перейдём теперь к другому и неудачному примеру.

3.7. Статистический прогноз. Проблема предсказания будущего всегда занимала человека. Греческие и римские полководцы никогда не начинали сражения, не посоветовавшись со жрецами, которые гадали по внутренностям жертвенных животных, прибегая, как бы мы сказали сейчас, к некоторой разновидности случайного эксперимента. Христианская религия отвергает такое гадание, но допускает пророческие сны и видения. Она косо смотрела на астрологию, подозревая в астрологах наклонность контакту с дьяволом, но формального обвинения такого вида не выдвигала. В средние века астрология продолжала существовать, не подвергаясь гонениям в целом как *научное направление*. Знаменитый полководец времён Тридцатилетней войны Валленштейн известен своим пристрастием к астрологии. В роли его астролога эпизодически выступал великий Кеплер [2]. В частности, в 1624 г. Валленштейн попросил Кеплера составить ему гороскоп. Кеплер довёл гороскоп до марта 1634 г., когда *Марс, Сатурн, Юпитер, Солнце, Венера и Меркурий расположатся в причудливом кресте*, что может привести к ужасным событиям. И вот, 25 февраля 1634 г. Валленштейн был убит по приказанию императора, не без оснований подозревавшего его в тайных переговорах с неприятелем.

Но этот блестящий пример прогноза не убеждает нас в ценности астрологии[3.9]. Мы склонны видеть здесь лишь удачное совпадение по точному образцу науки о благовониях (§ 1.1).

Теперь нам требуется либо статистический анализ всех астрологических прогнозов, удачных и неудачных, либо их какое-то научное обоснование. Поскольку ничего такого нет, мы больше не верим в астрологию. Но мы твёрдо верим в научные прогнозы, например, в расчёты будущих положений планет на основании закона всемирного тяготения. Здесь наша *вера* есть, собственно, уверенность, которая никогда не обманывает, хотя, как известно, общая теория относительности вносит в него поправку.

Могут ли статистические методы прогноза случайных процессов дать нам, если не столь твёрдую, но всё же разумную уверенность в предсказаниях будущего? Методы прогноза стационарных случайных процессов разработали независимо друг от друга Колмогоров и, несколько позже, Н. Винер.

В своих работах по теории турбулентности Колмогоров прямо говорит, что свои гипотезы о структуре этого явления считает весьма правдоподобными. Любопытно сопоставить с этим тот факт, что в работах по прогнозу случайных процессов он никак не намекает на возможность практических применений.

Указания на практическую важность теории прогнозов имеются в *Кибернетике* и других работах Винера. Так, в 1950 г. на Международном конгрессе математиков Винер [16] сказал, что его побуждала

Задача предсказания будущего положения самолёта на основе общих статистических сведений о способах его полёта и более конкретного знания прошлой траектории … Моя работа касается приборов, которые необходимы, чтобы реализовать теорию предсказания в автоматическом аппарате для стрельбы с упреждением по самолёту.

Известно, впрочем, что такой способ стрельбы с упреждением, который вытекает из работ Винера, не был реализован. Дело не в вычислительных трудностях, а прежде всего в том, что траекторию самолёта нельзя описать моделью стационарного случайного процесса. Возможно, в манёвре самолёта и есть статистическая составляющая, но как её выделить? В общем, манёвр настолько зависит от конкретных условий, что о статистической однородности всех траекторий полёта не может быть и речи. Можно пытаться выделить статистические элементы, но эта задача слишком сложна, и в условиях военных действий невыполнима.

И во всех других процессах, экономических, технологических, метеорологических и пр., мы обычно сталкиваемся с тем, что статистический элемент, если и присутствует, не охватывает всего явления в целом. Например, статистическому описанию может поддаваться лишь быстрая мелкомасштабная составляющая. Но и это ограничение лишь в исключительных случаях установлено научно, например, для микроструктуры турбулентности. Другой пример такого рода касается уходов частоты генератора колебаний за очень малое время, пока не успеют сказаться эффект мерцания и прочие технические причины этих уходов [7].

В огромном же большинстве случаев возможность статистического описания хотя бы какой-нибудь одной стороны изучаемого явления достоверно не установлена. Может помешать либо недостаточное количество экспериментального материала, либо непонимание необходимости сделать все статистические проверки, какие удастся придумать. В таких случаях статистический прогноз не более научен, чем прогноз на глаз, который в большинстве случаев и делается. Единственное преимущество статистического прогноза может заключаться в том, что он будет точнее глазомерного. Но, вообще говоря, надежды на это мало, потому что ошибка этого прогноза обычно велика. Например, если интересоваться прогнозом микронеоднородностей турбулентности (где статистическая модель установлена), то наилучший статистический прогноз значения $\xi(t + \tau)$ той или иной характеристики в момент $t + \tau$ по значениям $\xi(s)$, $s \le t$, почти не отличается от тривиального прогноза $\xi(t + \tau) = \xi(t)$. Следовательно, априорное преимущество статистического прогноза не очевидно и должно устанавливаться экспериментально.

Мы уже говорили об экспериментах Морана с прогнозом числа солнечных пятен, которые показали бесполезность статистического метода прогноза. Подойдём с этих же позиций к оценке недавно предложенного метода прогноза [11] и [12]. Метод состоит, собственно, из двух частей. Во-первых, полагается образовать разности от имеющегося ряда наблюдений и попытаться описать их моделью стационарного процесса, которая является комбинацией моделей авторегрессии и скользящего среднего. Если это не удаётся для первых разностей, надо вновь образовать разности от них и т. д.

Эта часть метода не вызывает особых возражений с той только оговоркой, что чем больше разностей мы образуем, тем больше потеряем информации о первоначальном процессе. В большинстве случаев мы сумеем описать разности достаточно высокого порядка, но как вернуться от них обратно?

На этот вопрос отвечает вторая часть метода. Надо сказать, что она математически некорректна. Например, рассматриваются суммы бесконечного числа одинаково распределённых случайных величин. Такие ряды всегда расходятся. Хотя в математике на расходящиеся ряды смотрят не так отрицательно, как прежде, потому что введение обобщённых функций позволило придать смысл многим из них, но это не относится к рядам указанного типа. Хуже всего то, что эти ряды формально подставляются в аппарат теории условных ожиданий. Таким путём можно получить всё, что угодно. Вторая часть метода видимо вообще неприменима к рядам наблюдений, которые описываются моделью тренда с ошибкой.

Но никаких статистических проверок, которые исключили бы эту модель, не делается. Вообще, все рекомендации направлены на то, чтобы смотреть лишь на корреляционные функции и забыть о самих наблюдениях, что в корне расходится с хорошей статистической традицией.

Таким образом, нет никаких гарантий научности предлагаемого метода прогноза и, в частности, того, что ошибка прогноза уложится в вычисляемые доверительные интервалы. Глазомерный метод совершенно равноправен, и вопрос лишь в том, какой метод даёт бОльшую ошибку.

Экспериментальный материал для решения этого вопроса крайне ограничен. В [11] и [12] имеется по существу единственный пример прогноза, см. наш рисунок 5. Сплошная ломаная – логарифмы выручки от продажи билетов на самолёты помесячно за 1949 – 1960 гг. Пунктирная линия – результат прогноза, сделанного по данным до июля 1957 г. Прямые сверху и снизу – результат эксперимента, состоявшего в том, что данные за август 1957 – 1960 гг. были прикрыты непрозрачным листком бумаги, а минимальные и максимальные точки за каждый год сглажены прямой линией. Продолжив их пунктиром, получаем глазомерный прогноз, который почти совпал с прогнозом в [11].

Максимальные точки соответствовали июлю или августу каждого года, минимальные – одному из зимних месяцев. Повторить эксперимент для других месяцев невозможно, потому что данные для них не читаются на рисунке, а таблица значений прогноза не приведена.

Удивительно, что описанный эксперимент не выходит на чертеже 9.2 [12], на котором изображено то же самое. Здесь прогноз идёт существенно лучше глазомерного и ближе к фактическим данным, хотя модель, её параметры и интервал прогноза одни и те же. Как же объяснить улучшение прогноза?

В общем, работы этой школы не выдерживают внимательного анализа. Используя выражение М. Булгакова, можно назвать такую статистику *статистикой облегчённого типа*[3.10]. Она рассчитана на всеобщую популярность, поскольку подаётся как универсально применимая и не требующая статистических проверок материала. Но она не обеспечивает надёжного результата.

Общий вывод состоит в том, что не следует особенно полагаться на статистические методы прогноза. Для применения этих методов следует прежде всего установить, действительно ли исследуемое явление описывается моделью случайного процесса. Лишь в этом случае можно полагаться на статистические методы.

Библиография автора

1. Белова Л. А., Мамиконянц Л. Г., Тутубалин В. Г. (1965), Вероятность аварийного пробоя изоляции обмоток статоров турбогенераторов и т. д. *Электричество*, № 4, с. 42 – 47. О статистической однородности корпусной изоляции статоров турбогенераторов. Там же, 1967, № 6, с. 40 – 46.

2. Белый Ю. А. (1971), *Иоганн Кеплер, 1571 – 1630*. М.

3. Дженкинс Г., Ваттс Д. (1971 – 1972), *Спектральный анализ и его приложения*. Вып. 1 – 2. М.

4. Ивахненко А. Г., Лапа В. Г. (1971), *Предсказание случайных процессов*. Киев.

5. Матерон Ж. (1968). *Основы прикладной геостатистики*. М.

6. Монин А. С., Яглом А. М. (1967), *Статистическая гидромеханика*, часть 2. М.

7. Рытов С. М. (1966, 1976), *Введение в статистическую радиофизику*. М.

8. Слуцкий Е. Е. (1960), *Избранные труды*. М.

9. Брошюра [i].
10. Хеннан Э. (1964), *Анализ временнЫх рядов.* М.
11. Box G. E. P., Jenkins G. M., Bacon D. W. (1967), *Models for Forecasting Seasonal and Non-Seasonal Time Series.* В книге *Spectral Analysis of Time Series*. New York, pp. 271 – 311.
12. Box G. E. P., Jenkins G. M. (1970), *Time Series Analysis, Forecasting and Control*. San Francisco.
13. Kendall M. G. (1946), *Contributions to the Study of Oscillatory Time Series*. Cambridge.
14. Кендалл М. Дж., Стьюарт А. (1976), *Многомерный статистический анализ и временнЫе ряды.* М.
15. Moran P. A. P. (1954), Some experiments on the prediction of sunspot numbers. *J. Roy. Stat. Soc.*, vol. B16, No. 1, pp. 112 – 117.
16. Wiener N. (1952), Comprehensive view of prediction theory. *Proc. Intern. Congress Mathematicians 1950*, vol. 2, pp. 308 – 321. Cambridge, Mass:
17. Yule G. U. (1927), On a method of investigating periodicities in disturbed series etc. *Phil. Trans. Roy. Soc.*, vol. A226, pp. 267 – 298.

Примечания

1.1. Именно такие мысли высказывал Кеплер (Sheynin 1974, § 8.1.1). Автор неоднократно употребляет выражение *истинное значение*, смысл которого истолковал Фурье [i, Прим. 4.3]. Автор никак не разъясняет его, мы же полагаем, что правильнее было бы упоминать *действительное* или *реальное* значение.

2.1. Лежандр первым опубликовал принцип, если не метод наименьших квадратов. Классической задачей того времени было установление наиболее подходящих параметров земного эллипсоида вращения по градусным измерениям. Термин нормальное распределение ввёл Пирсон. В начале §2.1 сглаживающая кривая появилась в результате применения МНКв. Уточним: в таких случаях, вопреки указанной выше классической задаче, по МНКв дополнительно устанавливается или проверяется та теория, на основании которой были составлены исходные уравниваемые уравнения (например. проверялось бы, что эллипсоид вращения действительно лучше всего описывает форму Земли).

2.2. О математической статистике, статистике и математике см. [i, Прим. 4.10].

2.3. Автор фактически рассуждает о задачах предварительного исследования данных, важного этапа работы теоретической (но не математической) статистики.

2.4. Ныне хорошо известна трёхмерная графика, раздел компьютерной графики. Её применяют, например, архитекторы, чтобы увидеть проектируемые ими сооружения.

2.5. Известно, что сглаживанием функций многочленами много и очень успешно занимался Чебышёв, но автор его даже не упоминает.

3.1. Столь же важен вопрос о том, как Бевeридж определял цены, существовавшие в столь давние времена.

3.2. См. [i, Прим. 2.5].

3.3. Об истории понятия функции см. Youshkevich (1977). Текст его статьи мы перевели с русского оригинала.

3.4. В связи со скользящим средним следует упомянуть его же, Слуцкого, статью (1935).

3.5. Кто же придумал пример со стрельбой горошинами, Юл или автор?

3.6. Введение ошибок измерения усложнило модель, но разве можно было забывать про них?

3.7. Ряд Бевериджа не прошёл проверки быть может ввиду обстоятельства, отмеченного в Прим. 3.1.

3.8. Обобщённое дифференцирование определяется для обобщённых функций.

3.9. У читателя неизбежно оставалось совершенно неверное представление о Кеплере как астрологе. См. § 2.2.4 нашей статьи (1974).

3.10. В *Мастере и Маргарите* Булгаков упомянул *Комиссию зрелищ и увеселений облегчённого типа*. Мы вспоминаем другое выражение того же автора из того же источника: *Осетрина* (статистика!) *второй свежести*.

Библиография публикатора

Винер Н. (1948, 1961, англ.), *Кибернетика*. М., 1983.

Слуцкий Е. Е. (1935), Эксперимент как метод исследования. *Ж. Геофизики*, т. 5, № 1, с. 18 – 38.

Sheynin O. (1974), Prehistory of the theory of probability. *Arch. Hist. Ex. Sci*., vol. 12, pp. 97 – 141. **S, G,** 30.

Youshkevich A. P. (1977), On the history of the notion of function. *Arch. Hist. Ex. Sci*., vol. 26.

III

В. Н. Тутубалин

Границы применимости
Вероятностно-статистические методы и их возможности

Москва, 1977

0. Введение

В 1972 и 1973 гг. вышли две брошюры автора [i, ii]. Первая была посвящена элементарным вероятностно-статистическим приёмам и методам, а вторая – несколько более сложным[0.1]. Основная мысль обеих брошюр состояла в том, что методы теории вероятностей, как, впрочем, и любой другой науки, нельзя применять без разбора к любым интересующим исследователя вопросам: существуют определённые границы их применимости.

Обе брошюры вызвали довольно многочисленные отклики, естественно, как положительные, так и отрицательные. Насколько известно автору, преобладали положительные, по крайней мере количественно. Хотя в чисто научных вопросах количественное преобладание на некотором непродолжительном отрезке времени может ничего и не означать, в делах издательских это не так. Возник вопрос о переиздании первой брошюры, однако автор постепенно пришёл к выводу, что за прошедшие пять лет её содержание в некотором специфическом смысле, см. ниже, устарело.

Дело, конечно, не в том, что пять лет назад вероятностные методы нельзя было применять в ситуации, когда изучаемое явление не обладает статистической устойчивостью, а теперь это стало возможным. Так могло бы случиться, если бы были разработаны новые методы, не требующие статистической устойчивости, но столь быстро наука не развивается. Но произошёл определённый, и доказуемый ссылками на публикации, сдвиг во взглядах на область применимости вероятностно-статистических методов. Если рассчитывать на перспективу, он делает почти ненужным доказательство столь простой вещи, как необходимость каких-то ограничений при использовании теории вероятностей.

Далее, веянием нашего времени несомненно является быстрое развитие конкретных статистических исследований. Вести такие исследования трудно: они требуют почти сверхчеловеческого терпения и настойчивости. Тем не менее, они возникают и ведутся. При одноразовом статистическом обследовании проверка статистической устойчивости практически невыполнима, а если выполнима, то только с отрицательным результатом. Однако, многократное, т. е. практически многолетнее статистическое исследование с проверкой выводов на всё новом материале сообщает вполне достаточную достоверность.

Точнее говоря, в процессе такого исследования всегда можно понять, что мы знаем достоверно, что с некоторым сомнением, а чего вообще не знаем. Поэтому в публикации, рассчитанной на широкого читателя, чрезвычайно важно показать, какова должна быть методика статистического исследования, которое приводит к достаточно достоверным выводам. Здесь нет общих математических результатов, так что такую вещь можно сделать только на примерах. Кое-что в этом смысле было опубликовано в двух предыдущих брошюрах автора, но сейчас подобную работу хотелось бы сделать полнее и лучше.

Наконец, цель публикации каждого автора состоит не только в том, чтобы поучать других, но и в том, чтобы самому научиться чему-то. В наших предыдущих брошюрах мы сделали несколько довольно крайних утверждений о практической бесполезности тех или иных приёмов. Например, были затронуты применение формулы Бейеса для медицинской диагностики, применение схемы Бернулли к судебным решениям, проблема *хвостов*, т. е. маловероятных значений случайной величины, многомерный статистический анализ, периодограммная техника и прогноз случайных процессов.

С такими крайними взглядами нетрудно спорить, по каждому конкретному вопросу достаточно указать хотя бы один пример успешного практического применения данного приёма. Естественно, что ни автор, ни кто-либо другой не знает полностью всей литературы ни по одному из перечисленных вопросов. Автор рассчитывал произвести в некотором роде выборку из бесчисленного количества исследований. Сторонники того или иного метода, применявшие его в своей работе, могли бы обидеться на крайнюю точку зрения автора и представить доказательства противного. Впрочем, насчёт применения схемы Бернулли к судебным приговорам в настоящее время никто, пожалуй, спорить не станет, это общепризнанный вздор[1]. Но все остальные вопросы вполне жизненны.

Таким образом, автор рассматривал публикацию своих утверждений не как окончательный итог, а как начало большой работы по лучшему выяснению фактического положения[0.2]. Предполагалось, что придётся иметь дело со сравнительно небольшим количеством фактического материала. Но не только в этом состоит существенное преимущество описанного способа *выборки* по сравнению со сплошным изучением публикаций. Ведь известно, что научные статьи обычно пишутся слишком кратко, так что их чтение превращается в расшифровку[0.3], в данном же случае все затруднительные вопросы можно было бы выяснить у авторов исследования.

Конечно, наряду с действительными научными возражениями мы получали и малосодержательные. Такими обычно являются сообщения о результатах исследования, если сообщающий о них знает об этих результатах лишь понаслышке. В таких случаях, поскольку никаких конкретных данных не приводится, всегда непонятно, достигнут ли успех исследования за счёт правильного применения теории вероятностей, или вопреки её неверному применению, что тоже не исключено. Например, успешная работа некоторой технической системы может быть достигнута как при правильной оценке характера случайных возмущений, так и потому, что конструктор просто не принял во внимание неверных вероятностных оценок, а руководствовался инженерным опытом.

Но желательный нами результат был в целом достигнут: возражения научного характера, хоть и в небольшом числе, были действительно получены. Они касались проблемы *хвостов*, прогноза случайных событий и возможностей периодограммного анализа. Что касается проблемы *хвостов* и прогноза, нам удалось указанным способом познакомиться с интересными и, вообще говоря, судя по первым результатам, обещающими исследованиями, ещё, однако, далеко не законченными. Поэтому наше утверждение о том, что до настоящего времени не известно ни одного достоверного практического применения соответствующих приёмов, снимать пока не пришлось. При всём своём отрицательном характере оно полезно тем, что подчёркивает важность практической работы в этих областях.

Замечательнее всего и научно неотразимым было возражение профессора В. А. Тимофеева об использовании периодограммной техники. Оказалось, что, в частности, они могут быть с успехом использованы при наладке систем автоматического регулирования для выделения характерных периодов возмущений и настройке регулятора на их подавление. Несмотря на то, что применяемая техника, строго говоря, не имеет вероятностного характера, мы сочли необходимым кратко описать приведённый Тимофеевым пример (§ 2.3).

Что касается многомерного анализа, то автор, занимаясь некоторыми вопросами медицинской статистики, познакомился с одним, по-видимому, многообещающим примером её применения (§ 2.2), и почти несомненно, что подобные её методы имеют не только медицинское, но и широкое техническое значение и для многообразных задач о надёжности оборудования. Однако, нужны большие усилия, чтобы превратить *почти несомненно* в просто *несомненно*.

Нам казалось полезным обсудить и проблему более общего характера: какие цели разумно ставить перед вероятностным исследованием? Естественно, что они не должны быть ни слишком частными (это неинтересно), ни слишком общими (их не достичь). В гл. 1 мы обсуждаем эту проблему на историческом материале.

Мы выражаем искреннюю благодарность редактору В. И. Ковалёву, по инициативе и постоянной помощи которого была написана эта брошюра.

Глава 1. Крайние мнения о теории вероятностей

1.1. *Странная метафизика* Лапласа[1.1]. В преподавании и в практической работе приходится встречаться, правда, всё реже и реже, с заблуждениями о реальных возможностях вероятностных методов. В нарочито грубой форме эти заблуждения можно выразить примерно так:

Рассмотрим некоторое событие. Мы явно не можем сказать, произойдёт ли оно или нет. Следовательно, оно случайно, а потому давайте изучать его методами теории вероятностей.

Если вы начинаете спорить, то могут привести несколько учебников, в которых, действительно, примерно это, но в менее грубой форме, и написано. И получается, что теория вероятностей есть особая наука, в которой из полного незнания можно сделать некие содержательные выводы.

Со многих точек зрения, исторической, психологической и пр., кажется интересным выяснить исторические корни этого заблуждения. Вообще говоря, изучать историю возникновения тех или иных взглядов, в частности, научных, чрезвычайно трудно, так как обычно требуется проанализировать огромное количество источников. Но теории вероятностей в некотором смысле повезло.

Примерно на рубеже XVIII – XIX столетий один из величайших учёных, Лаплас, суммировал и продвинул далеко вперёд как общую идеологию, так и конкретные результаты этой науки[1.2]. Будучи чрезвычайно трудолюбивым человеком, он оставил весьма подробное изложение своих взглядов и результатов (1812). Мы считаем допустимым ограничиться анализом одного этого источника, хотя, конечно, строгий историк науки не одобрит такой точки зрения. По-своему он будет прав: с точки зрения истории науки чрезвычайно важно изучать, например, эволюцию взглядов Лапласа и его взаимоотношения с другими учёными. Но наши цели в сущности узко прикладные.

В наш век бурного развития наукометрии[1.3] источнику нужно дать внешнее описание. Мы пользовались изданием книги Лапласа 1886 г., в котором воспроизведено её третье издание 1820 г. с дополнением 1825 г. В целом книга представляет собой огромный том объёмом примерно 58 печатных листов. Приятно заметить, что и в наше время лишь немногие монографии превосходят этот объём, а следовательно возможности человека в смысле написания огромных книг не изменилось существенно.

Книга делится на две совершенно разные по стилю части. Первая, *Опыт философии теории вероятностей*, является введением и сводкой содержания книги. Он подчинён непременному условию, чтобы в тексте не было формул. Например, формула

$$f(x)=\frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})$$

выражается в тексте словесной фразой, в которую ещё включены определения чисел π и *e*. Такие фразы, конечно, мало приспособлены для восприятия. Однако, *Опыт* содержит и массу материала, философского, общенаучного и прикладного характера, изложенного, на наш взгляд, самым прекрасным стилем. Не будь этот стиль столь прекрасен, нам, через полтораста лет, может быть и не пришлось бы бороться с попытками применять теорию вероятностей всюду и без разбора.

Опыт занимает примерно 12 печатных листов. Остальное содержание составляет собственно *Аналитическую теорию вероятностей*, в которой Лаплас пользуется анализом вволю и даже довольно странным для нас образом. Эта странность чрезвычайно затрудняет понимание второй части книги (в то время как понимание первой части затрудняется полным отсутствием аналитических формул, см. выше). Сегодня, видимо, трудно найти человека, который мог бы похвастаться тем, что полностью прочёл и понял собственно *Аналитическую теорию*[1.4].

Но *Опыт* читали многие, а попытки понять его математическую часть привели к созданию более строгих, и потому более понятных способов доказывать предельные теоремы теории вероятностей[1.5]. Нас будет интересовать исключительно *Опыт*.

Это – произведение довольно свободного жанра. Несомненно, что психология учёного вообще устроена так, что над полученными им научными результатами имеется своеобразная *надстройка*, состоящая из идей и эмоций общего порядка, возникающих на почве конкретных результатов и дающих веру, волю и энергию, необходимые для того, чтобы эти результаты получать. Обычно конкретные результаты публикуются, а *надстройка* остаётся достоянием узкого круга учеников и друзей. Но вот Лаплас опубликовал и то, и другое, чем оказал, на наш взгляд, неоценимую услугу.

Не стесняясь рамками чисто научной публикации, Лаплас ведёт в *Опыте* обширную полемику, в которой многим недурно досталось[1.6]: Паскалю за ряд мест из его *Мыслей*, несостоятельных с точки зрения оценки вероятностей свидетельств; автору *Нового органона* за то, что, рассуждая по индукции, он *доиндуцировался* до того, что считал Землю неподвижной, не признавая учения Коперника, и многим другим. Больше всех досталось великому Лейбницу в связи с суммированием ряда

$$\frac{1}{1+x}=1-x+x^2-x^3+\ldots \qquad (1.1.1)$$

в точке $x = 1$. Но прежде, чем перейти к собственно критике, Лаплас рассматривает следующий случай, который, хотя и слишком анекдотичен, чтобы быть верным, но зато хорошо характеризует отношение Лапласа к Лейбницу[1].

Рассматривая двоичную систему, Лейбниц полагал, что единица представляет бога, нуль – небытие. Высшее существо вытянуло из небытия все прочие существа, подобно тому, как в двоичной арифметике нуль есть нуль, но единицы с нулями изображают все числа. Эта идея так понравилась Лейбницу, что он через иезуита Гримальди просил сообщить её китайскому императору, большому любителю наук, в надежде, что, узнав такое, тот немедленно обратится в христианство.

Сообщив это, Лаплас [1814/1999, с 856 лев. столбец] говорит далее, что

Лейбниц, всё ещё руководившийся странной и очень развязной метафизикой,

рассуждал так: раз частные суммы ряда (1.1.1) при $x = 1$ попеременно обращаются то в нуль, то в единицу, то возьмём в качестве его суммы ожидание, т. е. 1/2. Как мы теперь знаем, этот способ суммирования далеко не глуп и может иногда применяться, но Лаплас спешит разгромить Лейбница, уже скомпрометированного предыдущим анекдотом.

Весьма примечательно, что теперь, через полтораста лет, мы с полным правом можем сказать о Лапласе те же самые слова: *руководился странной метафизикой*. Это никак не относится к конкретному научному содержанию его творчества, но в полной мере к его общим идеям, связанным с конкретным научным фундаментом. *Опыт* начинается словами (1814/1999, с. 835 лев.):

Я хочу изложить здесь без помощи анализа принципы и общие результаты теории вероятностей, изложенные в этой книге, применяя их к важнейшим вопросам жизни, большинство которых не что иное, как задачи теории вероятностей.

Какие же *важнейшие вопросы жизни* Лаплас имел в виду и как он связывал их с задачами теории вероятностей?

В теории вероятностей имеется центральная предельная теорема (ЦПТ), которая при определённых условиях устанавливает, что сумма

$$S_n = \xi_1 + \ldots + \xi_n$$

большого числа случайных слагаемых ξ_i имеет приблизительно нормальное распределение. Поэтому, если измерять отклонение случайной величины S_n от её ожидания ES_n в единицах $\sqrt{DS_n}$, мы будем получать значения случайной величины, подчиняющейся стандартному нормальному закону. Кратко это записывают в виде

$$\frac{S_n - ES_n}{\sqrt{DS_n}} \to N(0,1)\ .$$

Здесь N(0, 1) есть стандартный нормальный закон (с ожиданием 0 и дисперсией 1). Рассмотрим случай, когда $n \to \infty$, ожидания всех величин ξ_i одинаковы и равны *a*, дисперсии также одинаковы и равны σ^2, а сами случайные величины ξ_i независимы. Применяя известные правила, получаем

$$\mathrm{E}S_n = \sum_{i=1}^{n} \mathrm{E}\xi_i = na,\ DS_n = \sum_{i=1}^{n} D\xi_i = n\sigma^2,\ \sqrt{DS_n} = \sigma\sqrt{n}.$$

Для случайной величины, подчиняющейся нормальному закону N(0, 1), типичными являются значения порядка единицы по абсолютной величине. Например, вероятность значения, по абсолютной величине большего трёх, составляет примерно 0,003 (правило *трёх сигма*[1.7]). Мы получаем, что практически достоверно неравенство

$$\frac{|S_n - \mathrm{E}S_n|}{\sqrt{DS_n}} \le 3,\ \text{так что}\ |S_n - na| \le 3\sigma\sqrt{n}.$$

Допустим, что $a \neq 0$. Тогда типичным значением суммы S_n является *na*, а случайные отклонения не превосходят $3\sigma\sqrt{n}$, которая с ростом *n* растёт существенно медленнее, чем *na*. Итак, при большом *n*, *детерминированная составляющая na* преобладает по порядку величины над случайными колебаниями.

Таков чисто научный результат, известный по крайней мере в некоторых частных случаях, уже Лапласу. Посмотрим теперь, какая строится над ним *надстройка* философски-эмоционального характера.

Процитируем ещё одно место. Речь здесь идёт о том, что если некая империя слишком старается расшириться за счёт соседей, то соседи в конце концов соединяются в коалицию, которая и кладёт предел такому расширению, часто с немалым ущербом для зачинщика. Поэтому, пишет Лаплас (1814/1999, с. 842 лев.),

Важно, как для прочности, так и для счастья государств, не переходить этих границ, к которым они беспрестанно возвращаются от действия этих причин.

Этот вывод разумен, прекрасен и весьма актуален для посленаполеоновской Франции, но Лаплас прибавляет: *Это один из результатов исчисления вероятностей*.

Имеется в виду, что, как в ЦПТ детерминированная составляющая берёт верх над случайностью, так и в вопросах политики чему быть суждено, то и будет. Но надо ли обосновывать это последнее утверждение с помощью ЦПТ? Для современного читателя совершенно очевидно, что в данном случае речь может идти только о далёкой аналогии, характерной не для науки, а именно для метафизики, притом весьма странной.

Чуть далее со ссылкой на теорию вероятностей говорится о том, что если некое государство имеет далеко за морем колонию, то рано или поздно колония освободится, потому что она постоянно стремится к освобождению. В конце *Опыта* (там же, с. 863 лев.) говорится:

Ряд исторических событий обнаруживает постоянное действие великих принципов нравственности среди различных страстей и интересов, которые двигают во всех смыслах обществами.

Вывод отсюда таков: раз действие великих принципов морали постоянно, то, как учит ЦПТ, они всё равно возобладают над случайностью, а потому лучше их соблюдать, а не то худо будет. Такой вывод, хоть и весьма похвален, но с научной точки зрения явно ничем не лучше, чем обращение в христианство китайского императора, предпринятое Лейбницем.

Опыт кончается знаменитой фразой (там же):

Замечательно, что наука, начавшаяся с изучения игр, возвысилась до важнейших предметов человеческого знания.

Имеются в виду как раз указанные *политические приложения* теории вероятностей.

При всей странности метафизики в философско-эмоциональной сфере, Лаплас проявляет удивительную проницательность в конкретных приложениях теории вероятностей. Автор просматривал *Аналитическую теорию* со специальной целью – найти хотя бы одно неверное конкретное утверждение. Казалось бы, опираясь на полтораста лет развития науки со времён Лапласа, да ещё при такой странности его общефилософских взглядов, можно было бы найти и конкретные ошибки. У Лапласа есть ведь и конкретные задачи сомнительного характера о *вероятности судебных приговоров* и т. п.

Но оказалось, что очень не просто найти хотя бы одно неверное конкретное утверждение у Лапласа[1.8]. Великое множество данных им приложений можно разделить на три части: **1.** Явно и совершенно бесспорные приложения вроде вопроса о частичных переписях населения или об изменении частоты рождений мальчиков в Париже за счёт подкидышей. **2.** Приложения к обработке результатов астрономических наблюдений, судить о которых затруднительно, потому что требуется знакомство с обширным материалом. **3.** Явно сомнительные приложения, вроде задачи о вероятностях судебных приговоров, однако здесь Лаплас делает выводы столь осторожно, что никаких чисто научных ошибок просто не может быть[1.9].

О первой группе говорить нечего. Во второй группе можно заметить нечто поучительное. Лаплас (1814/1999, с. 844 прав.), в частности, приводит полученное им из обработки наблюдений отношение массы Юпитера к массе Солнца, равное 1:1071, предлагая держать пари миллион против одного, что относительная ошибка этого значения не превосходит одного процента. По современным данным (БСЭ, т. 40, 1957, с. 29) отношение масс больше указанного на два с лишним процента. Пари явно проиграно[1.10].

Однако, большой вопрос, в чём тут дело: в неправильной обработке наблюдений Лапласом или в систематической ошибке самих наблюдений, которая, как известно, не устраняется никакой статистической обработкой? Автору не удалось разрешить этого вопроса, хотя допустить наличие систематической ошибки весьма легко. Кроме того, нелишним будет заметить, что в самое недавнее время масса Луны была исправлена в третьем знаке, так что к точности современных цифр надо подходить с осторожностью. Если же склоняться к тому, что и Лаплас правильно обрабатывал наблюдения, и современные цифры верны, мы получим поучительный вывод о необходимости учитывать возможность систематических ошибок[1.11].

Интересно привести также точку зрения Лапласа на задачи о вероятностях судебных приговоров. В отличие, например, от Пуассона он (1814/1999, с. 862 лев.) не переоценивал реальности подобных задач:

Столько страстей, различных интересов и обстоятельств усложняют относящиеся к этим предметам вопросы, что они почти всегда неразрешимы.

Он рассматривал соответствующие математические задачи в сущности как модели в современном понимании этого слова, полагая, что выводы, полученные путём точных вычислений, всегда предпочтительнее, чем самые изысканные общие рассуждения. В качестве примера рассмотрим вопрос о желательном числе присяжных.

Вопрос об их оптимальном числе Лаплас не пытается решить. Единственным конкретным выводом (там же, с. 852 лев.) является осторожная рекомендация о том, что при 12 присяжных по-видимому желательно увеличить число голосов, необходимое для осуждения, до 9, вместо 8, установленных во Франции, потому что, исходя из решения модельных задач, число 8 не даёт достаточных гарантий от ошибочного осуждения[1.12].

Для дальнейшего важно отметить, что Лаплас охотно признавал существование задач, неразрешимых методами теории вероятностей. Заметим, что в наш век почти равнозначны две формулировки: **1.** Данная задача не является задачей той или иной области науки, и **2.** Данная задача является задачей этой области науки, но она неразрешима.

1.2. *Умозрительная* критика теории вероятностей[1.13]. Мы видим, что уже во времена Лапласа в теории вероятностей сложилось несколько противоречивое положение, когда конкретные результаты оказываются несравненно более скромными, чем широкие перспективы, о которых думал учёный. Необходимо подчеркнуть, что такое положение существует не в одной теории вероятностей. Например, широко распространено мнение, что физика занимается наиболее фундаментальными законами природы, из которых в принципе выводимы, или будут выведены в далёком будущем, законы других, например, биологических явлений. Биология также охотно говорит о необходимости, например, научиться управлять биосферой в целом.

Психология учёного, видимо, вообще устроена таким образом, что для занятия наукой совершенно необходима некая психологическая установка, которая придавала бы определённую гармоничность и всеобщность часто скромным и разрозненным конкретным результатам. В частности, передача от поколения к поколению неослабевающего интереса к научным занятиям вряд ли может осуществляться без выработки такой психологической установки. Например, если сказать школьнику, склонному выбрать физику своей будущей профессии:

Будешь всю жизнь сидеть у циклотрона и мерить неизвестно что,

то он, пожалуй, и не станет физиком. Если же сказать ему:

Будешь иметь возможность внести вклад в изучение самых фундаментальных законов природы,

результат будет иным.

Применение критерия практики для проверки истинности научных положений имеет, в первую очередь в фундаментальных науках, ту особенность, что практическая проверка часто выходит за рамки одного поколения. По одной этой причине передача интереса к науке от поколения к поколению жизненно важна.

С другой стороны, приведение этой общей установки в соответствие с реальными результатами также является важной задачей. Эта работа происходит во всех науках в различных обстоятельствах. В теории вероятностей накал страстей несколько выше, чем, скажем, просто в математике. Отчасти это может быть связано с тем, что стоявший у истоков теории вероятностей Лаплас[1.14], благодаря литературно-риторическим достоинствам своего труда, заложил слишком уж большое несоответствие между эмоционально-философской и конкретно-научной стороной.

Слишком широкие надежды общего порядка имеют тот эмоциональный недостаток, что немедленно превращаются в разочарования при соприкосновении с конкретной проблемой. В чисто научном плане их недостаток состоит в том, что при постановке новых задач исследователь не проявляет необходимой критичности. В результате силы и средства тратятся на напрасные поиски решения задач, безнадёжность которых очевидна при чуть более критическом взгляде.

Как бы то ни было, в науке выработаны определённые представления об области применения вероятностных методов. Фактически каждый учёный, занимавшийся каким-либо исследованием прикладного характера с применением теории вероятностей, вносил определённый вклад в эти представления. Но их отчётливая формулировка, притом блестящая и в чисто литературном смысле, принадлежит Мизесу. Он сам пытался также строить своеобразные математические основы теории вероятностей. Его попытки вызвали оживлённую критику, и в настоящее время общепризнанной аксиоматикой теории вероятностей, является аксиоматика Колмогорова, однако сама концепция практических применений в общем следует представлениям Мизеса[1.15].

Напомним в двух словах эту концепцию *статистической однородности* или *статистического ансамбля* (статистического коллектива). Для выяснения принципиальной стороны ограничимся простейшим случаем, когда в результате эксперимента может наступить или не наступить какое-то событие A. Обозначим через n_A число наступлений события в n экспериментах, предположительно повторенных в одинаковых условиях. Отношение n_A/n называется частотой наступления события A. Ещё до Мизеса статистики, например, Пуассон, занимавшийся вероятностями судебных решений, прекрасно понимали, что для применимости вероятностных методов к изучению события A должно выполняться свойство устойчивости частот: при увеличении числа n частота n_A/n должна испытывать всё меньшие колебания, приближаясь в каком-то смысле к пределу, который и принимается за вероятность $P(A)$ события A[1.16].

Мизес дополнил эти представления явной формулировкой ещё одного свойства, которое на интуитивном уровне также отлично сознавали статистики. Именно, если мы заранее, до проведения n опытов разобъём их на несколько совокупностей числом n_1, n_2, ..., то соответствующие частоты n_A/n_1, n_A/n_2, ... также должны быть близки друг другу, если только числа n_1, n_2, ... достаточно велики.

Экспериментальный материал должен разбиваться на части на основании прежней информации. Например, в одну совокупность мы решаем отнести опыты, проведённые летом, в другую – зимой. Частоты n_{A1}/n_1, n_{A2}/n_2, ... же становятся известными после проведения n опытов. Здесь n_{A1}, n_{A2}, … означают числа наступления события A в первой, второй, … из ранее выделенных совокупностей опытов. Практически вполне допустимо и полезно также разбиение на те или иные части уже имеющегося материала. Однако, при этом становится острым вопрос о вольной или невольной подгонке результатов под желательные выводы.

Указанное требование Мизеса важно. Пусть, например, событие A состоит в выпуске бракованной продукции, причём вероятность $P(A)$ испытывает, скажем, сезонные колебания. Запишем это в виде

$$P(A) = P_t(A) = p_0 + p_1\sin(\omega t + \varphi).$$

Здесь t момент наблюдения, p_0 и p_1 некоторые константы, подчинённые условию $P_t(A) \geq 0$. Пусть $t = 1, 2, ..., n$. Нетрудно доказать, что при независимых наблюдениях при разных t отношение n_A/n будет стремиться к p_0, если только $\omega \neq 2\pi$. В то же время разделение совокупности наблюдений по сезонам года, если мы действительно имеем дело с сезонными колебаниями, покажет нарушение требования Мизеса. Знание об их существовании может оказаться очень важным.

Показывая важность требования Мизеса, рассмотренный пример, однако, ставит весьма важный вопрос: допустим, что мы не знаем заранее, что существует сезонная периодичность. Как могли бы мы догадаться, что статистический материал нужно разбить по сезонам года? Существует ли какой-нибудь общий рецепт для выбора подлежащих исследованию групп наблюдений, или надо испытывать все возможные группы?

Можно ответить лишь следующее. Никакого общего рецепта нет, а испытывать все возможные группы бессмысленно, потому что одна из всех мыслимых групп будет всегда содержать все случаи наступления события *A*, а другая – все остальные случаи, так что равенство частот по этим группам будет нарушено так сильно, как вообще возможно. Выбор групп является делом интуиции исследователя, либо накопленной информации.

Хотелось бы поставить и такой вопрос: допустим, что требования Мизеса выполнены, достаточно ли этого для применимости теоретико-вероятностных выводов? Являются ли эти требования не только *необходимыми*, но и *достаточными*?

При столь большой общности речь могла идти лишь о каком-то варианте математической теоремы, которая устанавливала бы, допустим, что при условиях Мизеса верно какое-то утверждение, скажем, закон больших чисел. Но здесь опять вопрос, как выбирать группы? Если допускать все возможные, то такое требование противоречиво, потому что не может лежать в основе математического доказательства. Если не все возможные, то надо сказать какие, а это затруднительно[1.17].

Стоило нам даже не приступить к исследованию, а лишь чуть задуматься над простейшей задачей о наличии сезонных колебаний у вероятности выпуска брака, как немедленно приходим к выводу о том, что для решения данной конкретной задачи имеющихся в науке общих предписаний явно недостаточно. Автор не знает ни одного исключения из этого вывода. Мы ещё будем говорить о том, что из этого не следует невозможности решения вообще ни одной практической задачи, но сейчас заметим, что при всех недостатках изложенной концепции она всё-таки с полной ясностью устанавливает, что какие-то ограничения области применения вероятностных методов необходимы.

В чисто научном смысле этот вывод совсем не нов. Мы видели, сколь осторожен был Лаплас в тех применениях теории вероятностей, в которых такая осторожность требовалась. Пуассон, хотя и сделал неправильную в целом работу о вероятностях судебных приговоров, отлично понимал необходимость проверки ряда предпосылок на фактическом материале и кое-что проверил с отличным совпадением [i]. В общем, нет, пожалуй, исследователя, который не выбирал бы каким-то образом те задачи, в которых применение теории вероятностей может оказаться эффективным.

Речь может, следовательно, идти лишь о методике преподавания дела. ЧтО написать в учебнике для начинающего читателя, либо в статье для широкого обсуждения? Эти соображения вызывают к жизни особый сорт рассуждений, который мы и называем *умозрительной критикой теории вероятностей*.

Студент, начинающий изучать предмет, никаким конкретным материалом обычно ещё не владеет. Это относится не только к студентам чисто математических специальностей, для которых программа обучения вообще не предусматривает никакого конкретного материала**[1.18]**, но и к студентам прикладных специальностей, для которых теория вероятностей, вместе со всеми теоретическими дисциплинами сдвинута на младшие курсы. Если же в статье обсуждаются принципиальные вопросы, то она адресуется людям, у которых фактический материал по большей части есть, но у всех разный. Отсюда необходимость умозрительного обсуждения проблемы.

Подобные обсуждения строятся по единому принципу: раз необходимость ограничений в применениях теории вероятностей признаётся, то посмотрим, можем ли мы проверить выполнение этих ограничений на практике. Без труда выясняется, что ограничения вообще сформулированы слишком неопределённо, да и исчерпывающая проверка выводов невозможна. Примеры таких (?) рассуждений можно найти в работах автора [i] и (1972). В последнее время стали известны некоторые работы Ю. И. Алимова, опубликованные в ряде журналов, в частности в *Автоматике* издания АН УкрССР. Алимов пишет очень ярким стилем, и многое из его работ хотелось бы процитировать. Но приходится ограничиться одной цитатой (1974, с. 21, перевод с украинского любезно предоставил сам автор):

Правильности сопоставления <u>n измерений – n независимых случайных величин</u> на деле не угрожает никакая экспериментальная проверка. Однако, подобные сопоставления кладутся по традиции в основу многих разделов математической статистики, теории методов Монте-Карло, случайного поиска, рационализации эксперимента и ряда других внешне серьёзных дисциплин, которые, будучи недосягаемы для опытной проверки, так сказать, многозначительно присутствуют при разработке систем автоматического управления.

Алимов имеет в виду, что по одной выборке нельзя проверить ни независимости отдельных наблюдений, ни совпадения законов их распределения. Вообще, домысливание ансамбля многих возможных выборок к одной действительно наблюдаемой кажется ему недопустимым. Поэтому он предлагает упразднить основные понятия и приёмы математической статистики: доверительные границы, распределения выборочных характеристик, критерии согласия, состоятельность, несмещённость и эффективность оценок. В частности вопрос о неправильной оценке Лапласом доверительного интервала для массы Юпитера должен был бы решён просто, хотя, на наш взгляд, несколько жестоко: все доверительные интервалы инженеры применяют для того, чтобы уклониться от ответственности перед конкретным заказчиком. Классические формулировки ряда результатов, согласно Алимову (1974, с. 31 – 32), имеют смысл, существенно отличный от приписываемого им по традиции, и после прояснения их смысла становятся попросту неинтересными для прикладника.

Цитированная работа написана очень выразительно и прозрачно. Единственно, что осталось непонятным нам – почему интересам прикладника лучше соответствует мизесовская концепция или связанная с ней вторая аксиоматика Колмогорова, чем классическая теоретико-множественная аксиоматика? В любом случае предпосылки теории логически не могут быть проверены. Возможно, что понимать эту теорию надо так. Концепция, скажем, по Виллю – Постникову даёт другой умозрительный подход к прикладным задачам и при этом подходе традиционные приёмы математической статистики выглядят нелепо[1.19]. Следовательно, поскольку две умозрительные модели противоречат друг другу, по крайней мере одна из них весьма сомнительна. Однако, текст работы не даёт однозначных оснований для такого толкования.

Что касается взглядов Алимова на классическую теорию вероятностей, то по крайней мере в сопоставлении со взглядами Лапласа это действительно крайние взгляды. Мы с ними не согласны, но откладываем ответ до гл. 2. С другой стороны, мы легко можем себе представить фактический материал, знакомство с которым должно привести только к таким взглядам. Сейчас же заметим, что методические цели, которые только и может преследовать *умозрительная* критика теории вероятностей[1.20], в настоящее время представляется не то, чтобы достигнутыми, но такими, достижение которых в принципе обеспечено.

Планк (1966, с. 13) писал:

Я смог установить один, по моему мнению замечательный факт. Обычно новые научные истины побеждают не так, что их противников убеждают, и они признают свою неправоту, а большей частью так, что противники постепенно вымирают, а подрастающее поколение сразу усваивает истину.

Несомненно, хотя бы потому, что методика преподавания всегда следует за наукой, что и в преподавании в широком смысле, включая пропаганду тех или иных взглядов, происходит то же самое. Дело, конечно, не в том, что критические взгляды, выраженные, скажем, в работах автора или Алимова, меняют взгляды общества на проблемы теории вероятностей, Напротив, эти работы служат лишь для выражения изменившегося общественного мнения. Нужды нет, что в настоящее время, возможно, многие лекторы вузов продолжают повторять студентам:

Теория вероятностей изучает случайные события; случайным же называют такое событие, которое может произойти и не произойти.

Тем не менее, общественные взгляды изменились, и это отражается в новых учебниках. В недавно вышедшем учебнике А. А. Боровкова (1972) нет и следа *странной метафизики* Лапласа. В общем, нет сомнений, что простую истину о том, что в теории вероятностей, как и в других науках, необходимо тщательное сопоставление теории с реальностью, подрастающее поколение должно усваивать сразу.

Так в чём должно состоять такое сопоставление и как его искать? Алимов полагает, что оно в важнейших случаях невозможно. На самом деле научно аккуратные работы, в которых оно имеется, действительно довольно редки. Мы закончим эту главу обсуждением общего вопроса о том, на что здесь можно рассчитывать, а в следующей главе приведём некоторые конкретные результаты.

1.3. *Суеверие науки* и более реалистический взгляд. Предложение Алимова упразднить бОльшую часть математической статистики не является самым суровым из того, что можно сказать вообще о науке. У великого русского писателя Толстого (1910) есть целая глава под названием *Ложная наука*. Его основная мысль заключается в том, что *пустые науки*, такие, как математика, астрономия, физика вообще не дают ответа на основные моральные вопросы, вроде вопроса *Зачем я живу, и как мне жить*? К этому прибавляется, что содержание наук состоит из отдельных, мало связанных друг с другом обрывков знания, заинтересовавших неизвестно почему некую небольшую группу людей. Учёные же освободили себя от необходимого для жизни труда, под которым Толстой (*Полн. собр. соч*. т. 45, с. 296 и след.), прежде всего понимает труд крестьянина, и живут неразумной жизнью.

Чрезвычайно интересно посмотреть, что можно ответить в наше время на эти обвинения. В наше время, в связи с всевозрастающим могуществом науки, моральные вопросы обсуждаются особенно интенсивно, и обзор этих проблем можно найти, например, в статье А. В. Гулыги (1975). Что касается фрагментарности содержания естествознания, то в известной степени это справедливо: мы действительно имеем дело не с всеобъемлющей теорией, которая объясняла бы всю природу, исходя из единых принципов, а со многими теориями разных явлений (физика, химия биология и пр.) Многие чрезвычайно важные вещи ныне не поддаются научному анализу. Но значит ли это, что содержание науки сложилось случайно, в угоду прихотям отдельных людей? Постараемся показать, что это не только неверно, но и крайне несправедливо, как и утверждение, что учёные освободили себя от необходимого для жизни труда.

Но для того, чтобы это показать, позволим предложить для начала один пример. С. Н. Давиденков (1975) рассказывает об интересном феномене из жизни австралийских дикарей. Стремясь попасть домой до захода Солнца, дикарь прибегает к магическому ритуалу: он обламывает ветку дерева в надежде таким образом замедлить движение Солнца к закату. При всей нелепости такой надежды, с нашей точки зрения, само средство, по мнению Давиденкова, должно было иметь некоторый практический успех: дикарь уменьшает охвативший его страх и получает возможность действовать более целенаправленно и в самом деле попадает домой до захода Солнца. Отличие европейского учёного от австралийского дикаря состоит в том, что учёный нашёл бы в себе силы и интерес изучить, действительно ли практически эффективное средство имеет тот механизм действия, который ему приписывается (т. е. замедляет движение Солнца).

Этот пример ярко показывает нам важнейшую отличительную особенность науки – стремление к достоверному знанию. Единичный же успех может достигаться и другими средствами, например, с помощью магического ритуала, который, сняв стрессовое состояние с психики, помогает дикарю действовать не только рационально, но и в полной уверенности в достижении цели.

Теперь позволим себе некоторую, хоть и отдалённую, но полезную для уяснения проблемы ассоциацию. Сравним многовековое движение науки за достоверное знание с другим многовековым движением за освоение Севера и Востока страны, к примеру России. Как русскому крестьянину удавалось прижиться и построить деревни только в тех местах, где можно было заниматься хлебопашеством, т. е. практически по долинам рек, так и наука прижилась только там, где было возможно сравнительно достоверное знание. В результате мы видим на карте гроздья деревень вдоль рек, а в междуречьях практически нет населения. Так же мы видим, что некоторые области науки (небесная механика) разработаны хорошо и с большим избытком покрывают практические потребности, а множество не менее важных проблем, от прогноза погоды до предупреждения гриппа, мы ещё только учимся решать на научном уровне.

В другом месте Толстой (там же, с. 311) сравнивает естественные науки с удовольствиями, играми, катаньями, прогулками, с тем выводом, что удовольствия не должны мешать основному делу жизни. По-видимому, во времена Толстого учёные составляли столь незначительную прослойку населения, что великий писатель не имел случая почувствовать принципиально трудовой характер науки.

Короче говоря, естествознание – одна из многих областей человеческой деятельности, со всеми недостатками и достоинствами, которые из этого вытекают. Следовательно, например, критика теории вероятностей *умозрительного* сорта, наподобие изложенной в § 1.2, может преследовать лишь ограниченные цели. Действительно, она показывает, что предпосылки для применений теории вероятностей логически не могут быть проверены[1.21]. Но это относится к предпосылкам любой науки, и, хотя нелогичность бесспорно снижает достоверность знания, всё-таки во многих случаях выводы из теории вероятностей имеют вполне достаточную достоверность для признания их научными.

Как устанавливается практическая применимость и достоверность выводов из теории вероятностей обсуждали многие, в частности Лаплас в *Опыте*. Рассуждения Лапласа в этом месте довольно малосодержательны и сводятся к тому, что индукция ненадёжна. Установил же Бэкон с помощью индукции, что Земля неподвижна[1.22], а аналогия и того хуже. В рамках нашей метафоры (?) ответ предельно прост: практическая проверка производится трудом многих людей и многих поколений, которые вновь и вновь возвращаются к изучению данного вопроса.

Если на крестьянском поле лежало несколько крупных валунов, которые отцу приходилось объезжать, вспахивая поле, а у сына появилась возможность убрать эти валуны, то он так и сделает. Точно так же в науке не возбраняется подходить с новыми средствами к старым проблемам и либо подтверждать, либо опровергать предшествующие результаты. В статистике эта мысль преломляется так: при малом количестве данных ничего достоверного сказать и нельзя, но при продолжительном статистическом исследовании, когда поступает всё новый материал, в конце концов никаких сомнений не остаётся. Алимов прав, когда говорит, что по одной выборке вообще нельзя проверить, имеем ли мы дело с независимыми случайными величинами. Но положение резко меняется, если поступает ещё несколько выборок. В этом случае можно, в частности, проверить доверительные интервалы, вычисленные по одной выборке.

Автору приходилось сталкиваться с чувством неудовлетворения, которое вызывает у некоторых людей, приверженных к логичности мышлению, сама концепция статистической проверки гипотез. Эта концепция с самого начала фиксирует уровень значимости, т. е. некоторую ненулевую вероятность ошибочно отбросить справедливую гипотезу. Некоторым это кажется неприемлемым. Но процесс познания не состоит из единого акта проверки гипотезы, да и отвергая гипотезу, мы, к счастью, не выносим смертного приговора. Будут новые данные – будет и новая проверка гипотезы[1.23].

Точка зрения на науку как на некоторую область трудовой деятельности, которая не выше и не ниже любой другой области – промышленности, сельского хозяйства, рыболовства, вряд ли могла бы быть отвергнута Толстым. В поддержку этого предположения можно напомнить, что в той же книге (*Полн. собр. соч.*, т. 45, с. 300) имеется признание того, что в своей области познания материального мира наука сделала действительно большие успехи. Современное развитие не оставляет сомнения в том, что существует, в отличие от *ложной*, настоящая, истинная наука. Каковы практические выводы из изложенных соображений?

Если мы признаём науку одним из видов живой человеческой деятельности, то отсюда, с одной стороны, вытекает её незаконченность и фрагментарность в каждый момент времени, так как в живой деятельности всегда чего-нибудь не хватает, часто даже очень многого. С другой стороны, отсюда вытекает всеобщность: человечество всегда будет заниматься этим делом, стремясь расширить область достоверно известного.

В ряде областей применений математики к исследованию реальных явлений, в частности, и теории вероятностей, в настоящее время сложилось ненормальное положение, связанное с переоценкой практических возможностей этой науки. В таких случаях целесообразно подчёркивать неизбежную фрагментарность всех существующих применений. Слишком обширные замыслы в математике могут оказаться неисполнимыми, и их неизбежный провал создал бы для математики крайне нежелательную обстановку *подрыва доверия*, в которой наука не может нормально развиваться.

Так, одно время считалось, что если бы появилась возможность решать огромные задачи линейного программирования, охватывающие всю экономику страны, то на этой основе следовало бы перестроить всё экономическое планирование. Сейчас совершенно ясно, что такую задачу нельзя ни поставить, ни решить, хотя бы потому, что в такой глобальной постановке не имеют смысла понятия линейного программирования, такие, как *набор возможных технологий*. Но в результате оказалось, что исследования локальных задач, в которых линейное программирование может быть эффективным, развиты совершенно недостаточно.

Громоздкие, совершенно ни на что не пригодные концепции возникают, если глобальные задачи линейного программирования пытаются сочетать с вероятностным описанием возможной неопределённости. Здесь опять-таки могут иметь смысл лишь хорошо выделенные локальные задачи.

Вообще, в применениях теории вероятностей для описания неопределённых ситуаций чрезвычайно важно уметь достигнуть некоторого единства между степенью огрубления действительности, на которую ещё можно пойти в вероятностной модели, и количеством информации, которую надо извлечь из той же действительности для определения параметров модели. Положение здесь прекрасно описывается пословицей: *Хвост вытащишь – нос увязнет*.

Иными словами, модель, которая подробно и надлежаще описывает действительность, может требовать столь большой информации для определения своих параметров, что её технически невозможно собрать. Грубая же модель, требующая лишь небольшого количества статистической информации, может быть непригодной для описания действительности. В умении найти выход из этих затруднений и состоит главное требование к исследователю, который практически применяет теорию вероятностей.

Глава 2. Логичные и нелогичные применения теории вероятностей

Пять лет назад в популярной брошюре автор считал целесообразным излагать основы математического аппарата теории вероятностей. Но почти одновременно с выходом этой брошюры появилось достаточное число учебников по теории вероятностей, в которых математическая сторона была изложена с более чем достаточной полнотой. Но начинает обрисовываться традиция, полностью господствующая сейчас в преподавании математического анализа и ряда других математических дисциплин, которая резко отделяет математическое и прикладное содержание. В начале века учебники теории вероятностей были полны реальных примеров статистических данных, в новых же учебниках такие примеры исчезают.

Возможно, что происходит естественный процесс размежевания преподавания математической теории и её приложений. Например, если бы в учебнике математического анализа мы пожелали дать и приложения, то пришлось бы излагать механику, физику, теорию вероятностей и многое другое. Но остаётся фактом, что если приложения анализа естественно попадают в курсы и учебники механики и физики, то приложения теории вероятностей, исчезая из математических учебников, пока не попадают никуда. Стало быть, основные приёмы того, как надо работать с фактическими данными, как, в частности, судить о том, выполнены или не выполнены те или иные статистические предпосылки, вообще не попадают в учебники. Поэтому автору показалось уместным включить часть таких приёмов в эту брошюру. Они и образуют, так сказать, её *дидактическую* часть.

Все такие приёмы носят частный характер, и их естественно излагать на конкретных примерах. Но изложение нескольких примеров, которые по тем или иным причинам представляются автору важными, преследуют и другую, более широкую цель. Автор стремился показать, что, несмотря на возможную *логическую необоснованность*, вероятностное исследование может дать практически бесспорный результат.

Доверительные интервалы, критерии значимости и прочие статистические приёмы[2.1], против которых возражает, в частности, Алимов, прекрасно работают в этих примерах и позволяют сделать определённые практические выводы. Но, конечно, реальные применения теории вероятностей, как во времена Лапласа, так и теперь носят частный и конкретный характер. Отношение автора к всеобъемлющим глобальным построениям достаточно выражено в гл. 1.

2.1. *Об одном новом подтверждении законов Менделя*. Этот пункт посвящён изложению одной работы Колмогорова (1940), которая имела непосредственное отношение к проходившей тогда биологической дискуссии.

Сначала несложные теоретические сведения. Пусть проведён эксперимент, последовательные повторения которого образуют настоящий статистический ансамбль, а результатом его является значение некоторой случайной величины ξ. Результаты *n* экспериментов обозначают по традиции

$$x_1, ..., x_n \qquad (2.1.1)$$

(вместо $\xi_1,..., \xi_n$) и называют *эмпирической функцией распределения* $F_n(x)$:

$$F_n(x) = \frac{\text{число } x_i < x \text{ среди всех } x_1, ..., x_n}{n}. \qquad (2.1.2)$$

Эта функция изменяется скачками величины $1/n$ в точках (2.1.1); для простоты предположим, что среди этих чисел нет равных друг другу. Она зависит, следовательно, от случайных значений (2.1.1), реализованных в экспериментах, т. е. сама является случайной.

Кроме того, существует неслучайная (*теоретическая*) функция распределения

$$F(x) = P[\xi < x] = P[x_i < x] \qquad (2.1.3)$$

результата каждого эксперимента.

Колмогоров доказал, что при $n \to \infty$ величина

$$\lambda = \sup \sqrt{n} (\text{по } x) |\ F(x) - F_n(x)| \qquad (2.1.4)$$

имеет некоторое стандартное распределение вероятностей (распределение Колмогорова). Этот результат верен при единственном предположении, что функция $F(x)$ непрерывна. В настоящее время известно не только асимптотическое распределение (2.1.4), но и распределения (2.1.4) при всех значениях $n = 2, 3, \ldots$

Практический смысл эмпирической функции распределения состоит прежде всего в том, что её график наглядно представляет выборочные значения (2.1.1). В известном смысле $F_n(x)$ при достаточно больших n напоминает теоретическую функцию распределения $F(x)$. Удобно нарисовать обе функции распределения $F(x)$ и $F_n(x)$ в таком масштабе по оси ординат, чтобы $F(x)$ изображалась прямой линией.

Если $F(x)$ в точности не известна, но известно, что она лишь сдвигом и сжатием отличается от некоторой стандартной функции $F_0(x)$

$$F(x) = F_0(\frac{x - a}{\sigma}), \qquad (2.1.5)$$

где a и σ – параметры, то для построения масштаба употребляют произвольные значения a_0 и σ_0. И если, например, график функции (2.1.5) с параметрами a и δ_0 есть биссектриса координатного угла, то график функции $F(x)$ при $a \neq a_0$ или $\sigma \neq \sigma_0$ уже не будет биссектрисой координатного угла, но, очевидно, некоторой прямой[2.2].

Если параметры a и σ нам неизвестны, то можно сначала построить график эмпирической функции распределения, а затем сгладить его прямой на глаз и определить по этой прямой значения параметров a и σ. Эта процедура широко применяется в статистике, не столько ради получения значений параметров, сколько для того, чтобы поглядеть на данные экспериментов, представленные в форме, чрезвычайно удобной для глазомерной оценки. Вообще, если в статистике удаётся представить данные в форме, доступной зрительному восприятию, то возможности анализа необычайно расширяются.

Для случая нормального закона F_0 иногда бывает в продаже соответствующим образом разграфлённая (*нормальная*) бумага. Но вообще-то, если под руками имеется лист миллиметровки и таблица нужного закона распределения $F(x)$ или F_0 с параметрами a_0 и δ_0, то изготовление вероятностного масштаба занимает считанные минуты. Его самостоятельное изготовление имеет ещё то преимущество, существенное при малом объёме выборки n, что, зная этот объём, мы можем на оси ординат заранее разместить точки $1/n$, $2/n$, …, $(n-1)/n$, которые и являются значениями эмпирической функции распределения. Понятно, что точки 0 и 1 в вероятностном смысле всегда уходят в бесконечность.

Если пользоваться вероятностной бумагой промышленного изготовления, на которой ось ординат оцифрована в процентах, то значения размещённых точек приходится, при n некратном 10, интерполировать на глаз. Работать с вероятностной бумагой имеет смысл уже при самом небольшом числе наблюдений порядка даже единиц, так что глазомерная интерполяция не всегда удобна.

Кратко заметим, что существует и другой способ наглядного представления выборки, который называется гистограммой[2.3], Он состоит том, что данные группируют в нескольких интервалах и изображают число наблюдений в каждом интервале прямоугольниками, основания которых – эти интервалы, а площади пропорциональны числу попавших туда наблюдений. При большом числе наблюдений гистограмма похожа на плотность распределения случайной величины ξ.

Однако, гистограмма выразительна и почти не зависит от выбора интервалов группировки лишь при числе наблюдений порядка по крайней мере нескольких десятков. Автор решительно предпочитает во всех случаях применять эмпирическую функцию распределения, хотя гистограмма более распространена.

Критерий Колмогорова, основанный на статистике λ, см. (2.1.4), может применяться для проверки соответствия предполагаемого теоретического закона $F(x)$ данным наблюдений (2.1.1), представленным функцией (2.1.2). Однако, для этого теоретический закон должен быть точно известен. Распространённый прежде, но теперь постепенно уходящий в прошлое ошибкой является применение критерия Колмогорова для проверки гипотезы вида: *теоретический закон распределения является нормальным*. Дело в том, что нормальный закон определён с точностью до выбора параметров. В высказанной гипотезе эти параметры не указаны, и предполагается их определение по выборочным данным, естественно, с помощью оценок

$$a \sim \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i;\ \sigma^2 \sim\ s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Таким образом, вместо статистики (2.1.4), имеется в виду составление статистики

$$\sup(\text{по } x)\sqrt{n}\,|\,F_0[\frac{x-\overline{x}}{s}]-F_n(x)\,|\,. \qquad (2.1.6)$$

Здесь F_0 обозначает нормальный закон с параметрами 0 и 1.

Статистика (2.1.6) отличается от (2.1.4) тем, что вместо $F(x)$ в ней стоит величина F_0, которая зависит от (2.1.1) через $\overline{x}$ и s и потому случайна. Типичные значения статистики (2.1.6) существенно меньше значений статистики Колмогорова (2.1.4). Поэтому, применяя для статистики (2.1.6) распределение Колмогорова, мы завысим границы доверительной области, и будем принимать гипотезу нормальности чаще, чем следовало бы.

Итак, аккуратное применение критерия Колмогорова на практике как раз в наиболее элементарной, и потому чаще всего встречающейся ситуации невозможно. Он работает в тех случаях, когда имеется много выборок, к которым уже применены некие статистические критерии, но желательно с общей точки зрения взглянуть на многочисленные результаты использования этих критериев.

Перейдём теперь к сути вопроса о *подтверждении законов Менделя*. Имеется в виду следующая классическая ситуация. Некий признак имеет две аллели *A* (доминантная) и *a* (рецессивная). Сначала берутся две чистые линии с генотипами *AA* и *aa* и принудительно скрещиваются друг с другом[2.4]. Получается гибридная линия с генотипом *Aa*, фенотип которой соответствует признаку *A*. От гибридной линии, позволяя её особям свободно скрещиваться, получают второе поколение. Если принять гипотезу полной случайности сочетаний гамет, то вероятность генотипа *aa* равна 1/4. Поскольку лишь особи с генотипом *aa* проявляют в фенотипе признак *a*, то вероятность его проявления также равна 1/4. Итак, если всего получено n особей второго поколения, то число проявлений признака *a* в фенотипе может рассматриваться как число успехов μ в n испытаниях Бернулли с вероятностью успеха $p = 1/4$.

Таков простейший случай закона Менделя. К 1940 г. был накоплен обширный экспериментальный материал, из которого было видно, что такой простейший закон действительно наблюдается во многих случаях. Были обнаружены и существенные отклонения от него, которые могли быть связаны с различной выживаемостью особей различных генотипов и другими причинами.

Школа Лысенко стремилась доказать, что этот закон не действует. С этой целью были проделаны, в частности эксперименты Н. И. Ермолаевой (1939). Их отличие состояло в том, что материал рассматривался не суммарно по всем особям второго поколения, а по отдельным *семействам*. Что такое в данном контексте семейство, лучше объяснить на примере. В опытах с томатами семейством называются все растения второго поколения, выросшие в одном ящике. Каждый ящик засеивался семенами, взятыми из плодов *ровно одного* растения первого поколения. В общем, разбивка на семейства совершено естественна.

Когда были взяты данные по отдельным семействам, то оказалось, что частоты проявления генотипа *aa* довольно сильно колеблются, никогда не совпадая в точности с $p = 1/4$. Был сделан вывод, что законы Менделя не выполняются.

Однако, Колмогоров показал, что наиболее многочисленные серии опытов можно объяснить как раз на основе простейшей модели Менделя. Если для k семейств численностью $n_1, n_1, ..., n_k$ количества проявления рецессивной аллели суть $\mu_1, \mu_2, ..., \mu_k$, то из классической теоремы Муавра – Лапласа следует, что нормированные величины

$$\mu_i^* = \frac{\mu_i - n_i p}{\sqrt{n_i pq}},\ p = \frac{1}{4},\ q = 1 - p = \frac{3}{4} \qquad (2.1.7)$$

имеют приблизительно нормальное распределение $N(0, 1)$. Точность аппроксимации вполне достаточна при n_i порядка нескольких десятков. Таким образом, если смотреть на совокупность μ_i^*, как на выборку, и если модель Менделя верна, её теоретическое распределение есть стандартный нормальный закон.

Колмогоров рассматривал две наиболее многочисленные серии опытов Ермолаевой и соответственно две выборки вида (2.1.7) с 98 и 123 наблюдениями. […] Объективный способ оценки – вычисление статистики Колмогорова. Он получил $\lambda = 0{,}82$ и 0,75. Вероятности получить лучшее согласие (т. е. меньшее λ) были равны 0,49 and 0,37, так что полученные значения λ были вполне удовлетворительны.

Таким образом, чисто статистическое исследование превратило данные, казавшиеся опровержением законов Менделя, в их существенное подтверждение.

В работе Колмогорова, кроме исследований противников теории Менделя, упоминаются и работы его сторонников, в частности Енина (1939). Она не была подвергнута подробному анализу, однако указывалось, что согласие с законами Менделя оказалось слишком хорошим, отклонения частот от $p = 1/4$ по отдельным семьям меньше, чем должно было быть согласно основной модели испытаний Бернулли. Приведём поэтому основные результаты Енина.

Он рассматривал расщепление гибридов томатов по признаку: нормальные – картофелевидные листья. В зависимости от сроков высева в теплице семян гибридных растений (февраль или апрель) результаты были разбиты на две группы. […] Следует решить, как их статистически обработать.

В прикладной математической статистике применение каждого заданного статистического критерия объективно: после некоторых формализованных действий (в конечном счёте – арифметических операций) получается ответ: отвергается ли на заданном уровне значимости или не отвергается гипотеза, или же на каком предельном уровне значимости её можно было бы ещё принять.

Но вот какие критерии выбирать – вопрос в значительной степени субъективный. Ответ зависит от того, какие особенности материала кажутся подозрительными. Свои впечатления статистик более или менее подходяще превращает в статистические тесты. Общих правил нет, можно говорить только о примерах. В принципе любой результат наблюдений маловероятен, а при непрерывном законе распределения вероятностей, которым мы сейчас занимаемся, вероятность любого результата просто равна нулю.

Поэтому при любой обстановке можно подыскать критерий, который отбросит любую гипотезу. Чтобы не переусердствовать, следует допускать только такие критерии, которые в данной естественнонаучной задаче имеют содержательный смысл. Но, если не желать отбросить некоторую проверяемую гипотезу, то обычно можно подобрать такие критерии, которые её и не отбросят. Здесь речь идёт уже о добросовестности статистика.

Обратим сначала внимание на эмпирическую функцию для первой серии наблюдений. Вся она находится выше теоретической, и в общем неплохо сглаживается некоторой прямой, почти параллельной теоретической. Всё различие в некотором сдвиге влево. Мы отлично видим его, но не знаем, является ли он статистически значимым или нет, и нам следует применить критерий, основанный на выборочном среднем. Оно равно – 0.64, а его дисперсия равна $1/\sqrt{11} \approx 0,30;$ напомним, что проверяется гипотеза о стандартном нормальном распределении значений μ*. По модулю отклонение превосходит *два сигма*, т. е. высоко значимо. Строго говоря, первая серия опытов не подтверждает законов Менделя.

Но насколько велико отклонение от законов Менделя, которое мы должны признать здесь? Можно сразу сделать этот вопрос беспредметным, если объявить, что полученный результат заставляет сомневаться в наличии статистического ансамбля. Примерно то же самое: усомниться в независимости отдельных исходов эксперимента, но попробуем обойтись менее жестокими средствами.

Предположим, что каждое растение проявляет рецессивный признак независимо от других, притом статистически (?), но вероятность успеха p отлична от 1/4: $p = 1/4 + \Delta p$. Какой должна быть добавка Δp, чтобы объяснить сдвиг влево эмпирической функции распределения? Достаточно ли $\Delta p = -1/40$? Мы считали, что величина (2.1.7) при $p = p_0 = 1/4$ и $q = q_0 = 3/4$ имеет стандартное нормальное распределение. На самом деле, такое распределение будет иметь величина

$$\mu^* = \frac{[\mu - n(p_0 + \Delta p)]}{\sqrt{n(p_0 + \Delta p)(q_0 + \Delta q)}}.$$

Легко видеть, что при вычислении разности между μ_0^* и μ^* можно пренебречь изменением знаменателя. Мы получим

$$\mu_0^* \approx \mu^* + \frac{n\Delta p}{\sqrt{np_0q_0}}.$$

В различных опытах величины n разнятся друг от друга. Согласно нашей таблице $np = np_0 = n/4$ в основном превышают 50, так что $n \geq 200$. Поэтому систематическое смещение равно

$$\frac{n\Delta p}{\sqrt{np_0q_0}} = \frac{4\sqrt{n}}{\sqrt{3}}\Delta p \approx 30\Delta p$$

и $\Delta p = -1/40$ вполне объясняет систематическое смещение, равное – 0,64.

Кстати, глазомерная оценка этого смещения даёт – 0,58, что мало отличается от – 0,64, если учесть, что среднее квадратическое уклонение среднего арифметического равно $1/\sqrt{11} \approx 0{,}30$.

В настоящее время таблицы для распределения статистики

$$\lambda' = \sup\ (\text{по } x)\ |F(x) - F_n(x)| \qquad (2.1.8)$$

сосчитаны и для конечного n, см., например, Большев и Смирнов (1967). Для первой серии наблюдений (при $n = 11$) эта статистика равна 0,28. Это значение весьма умеренно значимо на уровне более 20%.

Применив этот критерий, мы, следовательно, не обязаны считать, что данные первой серии отвергают применимость законов Менделя. Но не так уж ясно, в каком выводе больше естественнонаучного смысла: то ли в том, что данные не согласуются с законами Менделя, но расхождение объяснимо незначительным изменением p (равным 10%), то ли в том, что данные, пожалуй, явно не противоречат законам Менделя.

Впрочем, у Енина есть некоторое объяснение возможной причины расхождения: при первом посеве в феврале растения в теплице страдали от недостатка тепла и света и значительная часть проросших семян погибла. Растения с рецессивным признаком вполне могли иметь несколько меньшую вероятность выживания. Это следовало бы проверить в специальном эксперименте. Окончательные данные первой серии можно рассматривать как некоторое скромное подтверждение законов Менделя.

Обратимся теперь ко второй серии. Соответствующая эмпирическая функция распределения плохо сглаживается прямой, – впрочем, по несколько субъективному мнению автора. Во всяком случае, разброс наблюдений существенно меньше, чем полагалось бы по стандартному нормальному закону. Чтобы подтвердить это статистическим критерием, проще всего вычислить сумму квадратов наблюдений. Она равна 2,85, тогда как при верной проверяемой гипотезе распределение для неё есть хи-квадрат с 14 степенями свободы. Следовательно, указанное значение практически невозможно, как видно из таблиц этого распределения.

Значение статистики (2.1.8) равно 0,33, что при $n = 14$ значимо примерно на 5% уровне.

Если смещение первой серии наблюдений получило некоторое разумное объяснение, то вторая серия имеет незначимое смещение (среднее выборочное равно – 0,21), но существенно меньшую, чем полагалось бы, дисперсию. Такое более точное, чем положено, выполнение законов Менделя вряд ли возможно. Наиболее вероятный статистический вывод состоит в том, что имела место вольная или невольная подтасовка результатов. О каком-то неблагополучии говорит и нарушение нормальности распределения (невозможность сгладить эмпирическую функцию распределения прямой), которое, правда, при данном числе наблюдений затруднительно обосновать статистическим критерием.

В общем, насколько удалось выяснить, экспериментальные данные приведены, видимо, не полностью. Итак, можно подтвердить законы Менделя, стремясь их опровергнуть, а можно усомниться в них, желая подтвердить, и всё это обнаруживается чисто статистическим исследованием.

На этом примере мы столкнулись с любопытным нарушением порядка, который устанавливается в математической статистике. Если действовать строго по науке, то статистические критерии должны выбираться ещё до проведения эксперимента. Лишь после него нужно подставить полученные данные в формулы и произнести приговор: отвергается гипотеза или нет. О её подтверждении вообще говорить запрещается.

На самом же деле критерии чаще устанавливаются, исходя из замеченных особенностей при глазомерном анализе материала[2.5]. Они служат для проверки того, являются ли эти особенности статистически значимыми или нет. Правда, установив в нашем случае, что полезны критерии, основанные на выборочном среднем и сумме квадратов выборочных значений, мы могли бы при анализе нового аналогичного материала строго следовать статистической науке[2.6], но тогда вновь появляющиеся особенности материала не были бы замечены.

Какого вида могли бы быть эти особенности? Например, на нашем рисунке показано некоторое ненулевое число наблюдений в области $\mu^* \le -3$. При верных законах Менделя вероятность попадания одного наблюдения в эту область равна 0,0014, а одного наблюдения из 100 примерно в 100 раз больше (в данном случае это почти точно). Итак, вероятность появления наблюдений на обоих чертежах в указанной области примерно равна $(0{,}14)^2 \approx 0{,}02$.

Мы получили отклонение от стандартного нормального закона, значимое на уровне ≈ 2%. Что же, законы Менделя всё-таки неверны? Ну, во-первых, мы подбирали критерий под известный материал, а во-вторых, совершенная разумность не состоит в догматическом следовании критериям значимости. Разумный ответ, видимо таков. Основная масса наблюдений прекрасно согласуется с законами Менделя, но возможно случаются резкие уклонения. Можно предположить, что некоторый недостаток числа проявлений рецессивного признака имеет какой-то биологический смысл (хотя бы существует связь с выживаемостью).

Впрочем, сказанное достаточно иллюстрирует ту простую мысль, что истина в науке устанавливается трудом ряда поколений и не всегда достижима при каждом отдельном акте исследования.

2.2. ***Никто не знает часа*** … Древнее изречение *Никто не знает часа своей смерти*[2.7] оказалось до некоторой степени поколебленным, конечно, в статистическом, а не в индивидуальном смысле, когда были составлены таблицы дожития и оказалось, что вероятность дожить до определённого возраста, хотя и подвержена колебаниям в зависимости от условий жизни, но не слишком значительным. Дальнейший шаг к индивидуальному прогнозу, основанный на технике многомерного статистического анализа, отчасти сделан, а отчасти делается, в настоящее время. Мы расскажем об одной из наиболее выдающихся работ в этой области, о так называемом Фремингемском исследовании (см., например, Truett et al 1967).

Как известно, сердечно-сосудистые заболевания представляют собой одну из центральных проблем современной медицины. Они проявляются различно, и одним из их наиболее частых проявлений является так называемая ишемическая болезнь сердца (ИБС). Согласно принятой нами классификации, к ИБС относятся случаи инфаркта миокарда, коронарной недостаточности, грудной жабы[2.8] и смерти от нарушений коронарного кровообращения. Мы хорошо знаем, что ИБС часто поражает людей, ещё находящихся в расцвете творческих сил, что делает проблему особенно острой.

Существуют некоторые довольно смутные представления о роли факторов современной индустриализированной жизни в развитии ИБС (малая физическая активность, нервно-эмоциональный стресс, нерациональное питание и пр.), а также о возможной роли генетических факторов. Вне всякого сомнения, эти представления чрезвычайно важны, но, разумеется, хотелось бы иметь, кроме общих, но недостаточно чётких и доказанных представлений, какие-то научные, т. е. достоверно доказанные сведения.

Они, может быть и не охватывая всей проблемы в целом, давали бы надёжную основу для каких-то практических выводов. В частности, важное значение имеют объективно устанавливаемые факторы риска, к которым, с одной стороны, относятся предвестники болезни, определяемые современными диагностическими средствами, например, изменения в электрокардиограмме (ЭКГ), а с другой – факторы жизни и поведения (возраст, курение, количество [*плохого*] холестерина в крови и пр.).

Поскольку речь идёт о некоторых точно определённых факторах, а не о смутно понимаемом *чрезмерном темпе современной жизни*, то в принципе возможно научное исследование их роли. Ввиду того, что возможные механизмы развития ИБС известны плохо, основным является здесь статистический метод исследования. Как обычно, в таких случаях основная надежда покоится на том, что недостаток сведений о природе явления (об ИБС) удастся возместить большим фактическим материалом. И ввиду того, что ИБС развивается постепенно в течение многих лет, желательно, чтобы исследование охватило не только большое число людей, но и длительный срок, по возможности, всю жизнь.

Само одноразовое обследование большой совокупности людей представляет большие трудности. Если же учесть, что в течение жизни люди обычно несколько раз переезжают с места на место, то понятно, что реальные трудности весьма велики. И следует понимать, что доля *случаев*, т. е. людей, у которых в конце концов развивается ИБС, довольно мала: обследуемая популяция состоит из *не-случаев*, т. е. из людей, у которых ИБС не развивается. Поэтому потеря из поля зрения исследователя *не-случая* сравнительно маловажна, но потеря хотя бы одного *случая* крайне нежелательна.

Но если мы допускаем потери, хотя бы ввиду переезда или отказа явиться для обследования, то не знаем, что мы потеряли: *случаи* или *не-случаи*. По этой причине потери из-за наблюдения стараются свести до минимума, тратя на это едва ли не бОльшую часть всех необходимых усилий.

Предметом Фремингемского исследования было практически всё население небольшого города Фремингема (США), имевшее в начале обследования возраст в пределах 30 – 62 лет. Исследование продолжается уже более двадцати лет, и в цитированной работе сообщаются его результаты за первые 12 лет. Наблюдались 2187 мужчин и 2669 женщин, у которых при первом обследовании не было ИБС. Из них за 12 лет развитие ИБС было обнаружено у 258 мужчин (11,8%) и у 129 женщин (4,8%), было давно известно, что женщины болеют ИБС реже мужчин. Рассматривалась связь между факторами риска, измеренными при первом обследовании, и вероятностью развития ИБС.

Факторов риска можно назвать довольно много, в исследовании учитывались семь: **1.** Возраст в годах. **2.** Содержание холестерина в сыворотке крови в *мг*/100 *мл*. **3.** Систолическое кровяное давление (*мм* рт. ст.). **4.** Относительный вес (вес тела, выраженный в процентах относительно среднего веса для соответствующего пола и возраста). **5.** Содержание гемоглобина (*г*/100 *мл*). **6.** Курение: 0 для некурящих, и далее для выкуривающих менее пачки сигарет в день – 1, одну пачку в день – 2, более пачки – 3. **7.** ЭКГ: 0 – нормальная, 1 – ненормальная.

При обработке наблюдений, результат которых зависит от многих факторов, возникает трудность совершенно общего характера, преодоление которой возможно и является главным результатом работы.

Возникновение ИБС, вообще говоря, малопонятным образом связано со значениями факторов риска. Когда их немного, допустим, один или два, обычно прибегают к классификации собранного материала по их возможным значениям. Для этого области значений каждого фактора разбивают на несколько интервалов (уровней), в простейшем случае на два, но это очень грубо, лучше выделить больше. Если значения каждого фактора разбиты на несколько уровней, то выделяют все появляющиеся при этом группы.

Частоты встречаемости ИБС в этих группах дадут оценки соответствующих вероятностей, что и будет надлежащим описанием материала (с некоторым огрублением ввиду огрубления значений факторов риска).

Например, область значений содержания холестерина можно разбить на четыре интервала: до 190, от 190 до 219, от 220 до 249, и не менее 250 *мг*/100 *мл*. Также и область систолического давления: не более 129, от 130 до 139, от 140 до 149, и не менее 150 *мм* рт. ст.

Далее мы устанавливаем двумерную классификацию: составляем группу, у которых холестерина менее 190 и одновременно давление не более 129 и определяем частоту возникновения ИБС в ней и т. д. Всего окажется 16 групп, и частоты возникновения ИБС вычисляются уже не по всем 4856 наблюдениям, а по количеству наблюдений в среднем, в 16 раз меньшему. Весьма вероятно, что мы не сочтём возможным объединять мужчин и женщин, и тогда число наблюдений уменьшится ещё, грубо говоря, вдвое.

В общем, для каждой частоты получается уже скромное число наблюдений порядка сотни. Но что получится, если прибавить ещё три группы по возрасту? И ещё 4 по курению? И не забудем про 2 группы по ЭКГ и т. д. И в результате получим классификацию, каждая группа которой состоит в лучшем случае из одного наблюдения, и не исключены случаи пустых групп. Следовательно, мы не сумеем определить по частотам никаких вероятностей. Например, при 10 факторах, рассматриваемых на трёх уровнях каждый, многомерная классификация будет содержать 59 049 групп. Каков же должен быть объём материала, чтобы в каждой группе оказалось хотя бы несколько десятков наблюдений?

Эта же трудность встречается и во многих технических задачах о надёжности работы оборудования, для которого установлено несколько видов испытаний. Если эти испытания дали результаты

$$x_1, ..., x_k, \qquad (2.2.1)$$

то хотелось бы вывести вероятность безотказной работы как функцию $p(x_1, ..., x_k)$, но бессмысленно пытаться сделать это путём многомерной классификации.

Посмотрим, как эта задача решалась во Фремингемском исследовании. В их решении имеется одна бесспорная часть, насколько можно сейчас судить, а другая часть совершенно не логична. Это не значит, что она по существу неверна, но возможно, что она нуждается в каких-то в уточнениях. Бесспорную часть можно изложить так.

В общем-то, из функций нескольких переменных единственно хорошо изученной является линейная. Существует целая наука, линейная алгебра, которая изучает линейную функцию многих переменных (и отчасти квадратичную). Поэтому хорошо бы изобразить неизвестную вероятность $p(x_1, ..., x_k)$ в виде линейной функции. Это, однако, явно невозможно, ибо вероятность изменяется от 0 до 1, а линейная функция не ограничена. Поэтому необходимо дальнейшее усложнение: положим, что

$$p(x_1, ..., x_k) = f(a_0 + a_1x_1 + ... + a_kx_k).$$

Аргумент здесь – линейная функция, а $f(y)$ – функция одного переменного, изменяющаяся в пределах между 0 и 1. Многие соображения простоты указывают, что удобнее всего принять так называемую логистическую функцию

$$f(y) = \frac{1}{1+e^{-y}}.$$

Окончательно, меняя обозначения, чтобы привести их в соответствие с обозначениями цитируемой работы, мы получаем основную гипотезу вида

$$p(x_1,...,x_k) = \frac{1}{1-\exp[-\alpha-\sum_{i=1}^{k}\beta_i x_i]}. \qquad (2.2.2)$$

Правая часть называется *многомерной логистической функцией*. Разумеется, мы не доказали, что искомая вероятность *p должна* выражаться функцией (2.2.2), но пришли к ней без особых логических несообразностей.

Теперь возникает вопрос об оценке параметров этой функции по результатам наблюдений. Вот здесь авторы работы совершенно сознательно допускают логическое противоречие. Они предполагают модель многомерного нормального распределения для результатов обследования (2.2.1). Это явно невозможно, поскольку среди семи факторов два (ЭКГ и курение) вообще измеряются в дискретных единицах, и о нормальности, говоря формально, нет и речи. И довольно странно считать возраст нормально распределённым. В отличие от мелких нелогичностей, допущенных в § 2.1, вроде подбора статистического критерия под уже имеющиеся данные, здесь речь идёт о крупной нелогичности, которую может оправдать только полученный результат (*победителей не судят*).

Точнее говоря, основная гипотеза исследования состоит в том, что имеются две многомерные нормальные совокупности. Одну составляют наблюдения факторов риска у лиц, которые в последующие 12 лет не заболели ИБС, а другую – те же наблюдения над заболевшими впоследствии. Ставится дискриминантная задача о различении этих совокупностей. Классическое предположение дискриминантного анализа состоит в равенстве ковариационных матриц двух совокупностей. Это предположение приводит (довольно примечательный факт!) к тому же выражению (2.2.2) для вероятности заболевания лица с факторами риска (2.2.1)[2.9], к которому мы пришли из соображений простоты. Сейчас эта вероятность понимается как апостериорная вероятность заболевания при условии, что измерения факторов риска дали (2.2.1). Но на этот раз получается и способ оценки параметров функции (2.2.2), нелогичный в той же мере, в какой нелогично предположение нормальности.

Наши рассуждения, которые вначале привели нас к выражению (2.2.2), привели бы к совершенно иному вычислительно более сложному способу оценки этих параметров, к способу наибольшего правдоподобия. Этот способ тоже может быть реализован, и считается, что применительно к данному материалу оба дают весьма близкие результаты.

Интересно, однако, узнать, какие практические выводы получены в цитируемом исследовании. Вначале, оценив каким-то способом значение параметров, мы можем по формуле (2.2.2), в которую вместо параметров подставлены их оценки, узнать для каждого обследованного лица приближённое значение вероятности $\hat{p}$ возникновения у него ИБС в последующие 12 лет.

Наивысшие вероятности по возрастным группам таковы.

Для мужчин в возрасте 30 – 39, 40 – 49 лет и 50 – 62 года 0,986, 0,742 и 0,770 (заболевание наступило, наступило и не (?) наступило). Для женщин в возрасте 30 – 49 лет и 50 – 62 года – 0,838 и 0,773 (заболевание наступило).

Эти результаты в какой-то мере опровергают классическое изречение в заголовке этого § 2.2. Но надо иметь в виду, что формально речь идёт о *предсказании* уже совершившихся событий. Такое же предсказание будущего возможно лишь, если предположить, что в другом месте или в другое время коэффициенты функции (2.2.2) будут примерно теми же, какими их оценили авторы. Это предположение вероятно ещё не проверено.

Далее, упорядочив набор значений $\hat{p}$ для всех обследованных, можно разбить их на несколько равных по численности групп (например, по 10 человек) таких, что в первую группу входят лица с наименьшими значениями $\hat{p}$, во вторую – с бОльшими, в третью – с ещё бОльшими и т. д. Границы этих групп, если их 10, называются *децилями*. Если найденная линейная комбинация факторов риска не имела никакого отношения к ИБС, то во всех группах число случаев ИБС было бы примерно равным. На самом деле картина принципиально иная. Приведём для примера таблицу 3 из исследования авторов. *Ожидаемое* в этой таблице означает число ожидаемых случаев ИБС, полученное суммированием вероятностей $\hat{p}$ по всем членам соответствующей группы.

В этой таблице прежде всего поражает многократное, в несколько десятков раз, отличие процентов заболеваемости в группах высокого и низкого рисков. Во-вторых, несмотря на явную необычность распределений, результаты, полученные с помощью нормальной модели, неплохо согласуются с фактическими данными.

Таким образом, выделение групп с повышенной опасностью возникновения ИБС возможно при помощи самого простейшего клинического обследования, дающего перечисленные выше факторы риска[2.10]. Тот же самый вывод можно сделать на основании данных, представленных по отдельным возрастным группам.

Не следует думать, однако, что эти результаты годятся для индивидуального прогноза. Такой прогноз можно успешно делать только в случае очень большого или очень малого индивидуального риска, но для всей совокупности лиц результаты были бы плохими. Это связано с тем, что ИБС встречается всё-таки редко (11,8% в среднем за 12 лет[2.11]). Действительно, основываясь на значениях $\hat{p}$, для индивидуального прогноза мы можем только предсказывать заболевание у лиц с достаточно высоким $\hat{p}$ и его отсутствие у остальных. Взяв в качестве критического значения $\hat{p}$ границу группы наивысшего риска, и предсказав заболевание, мы ошиблись бы в 100 – 37.5 = 62.5% случаях. С другой стороны, в 258 – 82 – 176 случаях (в 68.5% всех случаев) ИБС наступила бы несмотря на наше обещание. Задача индивидуального прогноза далека от решения.

Посмотрим теперь на оценки коэффициентов в формуле (2.2.2) и на те выводы, которые можно из них сделать. Эти оценки различны для разных возрастных групп, а также для женщин и мужчин. Приведём результаты для группы мужчин всех возрастов. Оценка $\hat{\alpha}$ = – 10,8986[2.12]. Остальные оценки сведём в табл. 4. Из неё видно, что коэффициенты двух из семи факторов, относительного веса и гемоглобина, сравнительно мало превышают по абсолютной величине свои среднеквадратические ошибки. Эти факторы следует признать менее влияющие на ИБС, чем остальные пять.

Из этих пяти возраст не подлежит никакому управлению, а действие остальных факторов удобно соотносить с влиянием возраста. Например, 1 пачка сигарет в день есть 2 балла, чему соответствует приращение линейной функции, равное 0,7220, а это примерно соответствует 10 годам возраста. Иначе говоря, тот, кто выкуривает в день пачку сигарет, ускоряет в статистическом смысле наступление инфаркта миокарда (?) на 10 лет. Эта пропорция остаётся примерно одинаковой по различным возрастным группам мужчин. Для женщин данные другие: вред курения выражен гораздо слабее. Не вполне ясна причина этого – природа вещей или малое число курящих женщин во Фремингеме (из 2669 женщин 1562 вовсе не курили, пачку сигарет в день выкуривали лишь 301), что привело к неполноте статистики.

При помощи табл. 4 влияние холестерина и кровяного давления сопоставлять с влиянием возраста неудобно, потому что коэффициенты линейной комбинации любых размерных величин также размерны, в нашем же случае требуется, чтобы линейная комбинация была безразмерной. Сопоставление использованного нами вида даёт, например, такой результат как 7*мг* % (?) холестерина эквивалентны одному году жизни, но вот 7*мг* % холестерина это много или мало?

Чтобы ответить на этот вопрос, надо знать, как велики бывают колебания в содержании холестерина. Иначе говоря, можно разделить содержание холестерина на его среднеквадратическое отклонение и тем самым выразить его в безразмерных величинах. Если проделать такую операцию со всеми факторами риска, то сравнение коэффициентов при них даст такое расположение факторов в порядке убывания важности: возраст, холестерин, курение, кровяное давление, ненормальность ЭКГ, а вес и гемоглобин влияют меньше. Некоторая условность такой нормировки, которая имеет лишь чисто статистический смысл, должна быть понятной.

Здесь приходит в голову постановка вопроса вроде следующего. Если 7 *мг*% холестерина эквивалентно такому же числу лет, как и 4,5 *мм* рт. ст. кровяного давления, то что легче: снизить на 7 *мг*% содержание холестерина или на 4,5 *мм* кровяное давление? Вместо этого вопроса ставится вопрос о широте разброса, но это не то же самое. Однако, ставя вопрос о сравнительной лёгкости влияния на холестерин и кровяное давление, мы по всей вероятности заходим в научный тупик: скорее всего, эти вещи просто количественно несравнимы.

В общем, господствующий сейчас стиль мышления в отношении оптимальных задач по каждому случаю жизни обычно очень скоро приводит в тупик. Кроме того, цитируемая работа относится к совокупности людей, на которую никто не пытался влиять фармакологическими средствами. Совершенно неясно, сохранятся ли численные характеристики такой совокупности, если попытаться на неё повлиять. Скорее всего, не сохранится, а тогда всякое сопоставление холестерина с давлением теряет смысл.

Таким образом, этой работе никак нельзя приписывать возможность указания желательного пути влияния на факторы риска. По нашему мнению, чисто лечебного прикладного значения эта работа совсем не имеет. Но есть определённый круг исследований, для которых её роль должна быть очень велика.

Мы имеем в виду испытания различных лекарственных препаратов.

Единственным путём для получения здесь достоверных результатов является статистическое исследование. В простейшем случае случайно формируются две совокупности лиц, из которых одна (экспериментальная) получает исследуемый препарат, а другая (контрольная) – плацебо, т. е. безвредное вещество, на вид похожее на этот препарат. Кроме сравнения результатов по этим группам, не возбраняется сравнивать их с общими статистическими сведениями (со средней заболеваемостью ИБС в городе, стране и пр.). К сожалению, в современной кардиологии часто обнаруживается *эффект плацебо*: результаты по обеим группам практически совпадают, тогда как средние результаты по большей совокупности гораздо хуже. Этот факт довольно правдоподобно объясняется тем, что кроме лекарств врач помогает пациенту и другими способами, например, советами о рациональном образе жизни. (В современных экспериментах лечащий врач и сам не знает, что получает пациент – лекарство или плацебо.) И вот на фоне общего лечения квалифицированным врачом эффект химиотерапии, если он есть, совершенно не заметен. С другой стороны, для подавляющей части населения, либо не обращающейся к врачу, либо лечащейся у врачей более низкой квалификации, результаты (?) оказываются намного хуже.

Эффект плацебо лишает возможности составлять представление о реальной пользе фармакологических средств. Можно предположить, что дело здесь в том, что частота ИБС не столь уже высока. Если в крайнем случае допустить, что частота заболеваний некоторой болезнью равна 1%, а некое профилактическое средство снижает её вдвое, до 0,5%, то для сколько-нибудь надёжного установления эффективности этого влияния надо испытать препарат хотя бы на тысяче пациентах (вместе с контрольной группой, на двух тысячах). При этом в контрольной группе число заболевших μ_1 будет починяться закону Пуассона с параметром 10, а в экспериментальной группе соответствующее число μ_2 будет иметь то же распределение с параметром 5. Вероятность неверного результата $\mu_2 \geq \mu_1$ равна, грубо говоря, $\Phi(-5/\sqrt{15}) = \Phi(-1,29) \approx 0,10$, т. е. не очень мала[2.13]. В то же время, если бы удалось отобрать для эксперимента группу лиц с вероятностью заболевания 20%, то, снова при снижении вероятности заболевания вдвое, до 10%, сто пациентов удовлетворили бы нас гораздо лучше.

Фремингэмское исследование как раз и показывает, что в принципе отбор лиц с многократно увеличенным риском заболевания достижим при помощи простого клинического исследования. Впрочем, мы предполагали, что относительная эффективность препарата одинакова во всех группах риска. Если это предположение кажется необоснованным, то может быть неразумно было бы ограничить эксперимент группой наивысшего риска. Но в любом случае по окончании эксперимента крайне необходимо разбить испытуемых по группам риска, чтобы при сравнении групп примерно одинакового риска посмотреть, не обнаружится ли что-нибудь более содержательное, чем эффект плацебо. Иными словами, аппарат Фремингемского исследования должен стать составной частью при обработке по крайней мере многих лекарственных препаратов.

Естественно, что сказанное относится и ко многим техническим средствам, призванным повысить надёжность работы оборудования. Ввиду такой важности методики Фремингемского исследования возникает вопрос об обоснованности его результатов. Это исследование является примером не только получения важных результатов, но и, видимо, всестороннего обсуждения самими авторами возможных сомнений и возражений.

Они опубликовали не только положительные, но и некоторые отрицательные результаты. В частности, все численные оценки были основаны на 12-летнем наблюдении определённой популяции, но применимы ли они к другим популяциям? А если спросить совсем остро: не является ли наблюдаемое деление на группы различного риска простым артефактом, связанным с подбором довольно большого числа параметров? Известно ведь. что при попытках разделения совокупностей по большому числу признаков раз навсегда собранный материал можно неплохо разделить, но что при добавлении новых наблюдений старые формулы перестают действовать.

Авторы попытались применить полученное выражение для риска к выделению тех совокупностей, на которых должны были произойти новые, после 12-летнего срока, случаи ИБС. Опыт вполне удался[2.14] для мужчин в возрасте 30 – 39 лет и для женщин в возрасте 30 – 49 лет: в группе высокого риска новых случаев 8 у мужчин и 10 у женщин, в группе низкого риска, соответственно. 2 и 5. Но в других возрастных группах опыт оказался весьма неудачным. Возможно, что для старших групп такое простое клиническое обследование более, чем 12-летней давности совсем уж не показательно.

Применимость конкретных числовых значений констант в формуле (2.2.2) к другим совокупностям можно установить только экспериментально. Автору неизвестно, было ли это сделано. Обычные оценки, основанные на модели двух нормальных совокупностей, не допускают здесь возможности артефакта.

Таким образом, если при анализе одномерных выборок (§ 2.1) можно было непосредственно убедиться в истинности анализа, просто посмотрев на данные, представленные в удобной для зрительного восприятия форме, то в многомерном анализе так не бывает. Это очень затрудняет статистические исследования в многомерном пространстве.

2.3. Периодограмма для подавления колебаний. Публикация двух брошюр доставила автору удовольствие весьма примечательного знакомства с В. А. Тимофеевым, профессором широко известного Ленинградского электротехнического института (ЛЭТИ). Ныне Владимира Андреевича, к сожалению. уже нет в живых, он умер 5 апреля 1975 г., но осталось несколько его книг (например, 1960; 1973; 1975), описывающих мало кому сейчас известный мир эффективных математических приёмов.

В связи с широким развитием машинной математики, которая чрезвычайно расширила сферу практически выполнимых расчётов, внимание к простым, в частности к графическим методам вычислений[2.15] в наше время ослабло. Например, только из книг Тимофеева автор узнал, что такое ортогон Лилля (это графическая процедура для вычисления значения многочлена, пригодная и для отыскания корня). Конечно, машина выполняет эту операцию несравненно быстрее, но в тех случаях, когда многочлен появляется в технической проблеме, и есть возможность подходяще подбирать его коэффициенты, графическое решение предпочтительнее. Многие подобные приёмы представляют собой нетривиальные изобретения, подобные изобретениям машин или механизмов, до которых почти невозможно додуматься самому. Эти изобретения делались на протяжении веков, но в настоящее время интерес к ним, к сожалению, значительно упал. Метафора, сравнивающая развитие науки с освоением новой территории (§ 1.3), и здесь сравнительно точно рисует картину: упал спрос на определённые продукты, и посёлки, жившие их производством, заброшены. В науке, как и в прочих, не менее серьёзных вещах, многое зависит от прихоти моды.

Интересно привести мнение Тимофеева о критике теории вероятностей того сорта, который здесь назван у*мозрительным* (§ 1.2). В ней речь идёт всегда о том, что мы не в состоянии логически доказать выполнимость предпосылок теории. Тимофеев заметил, что такие рассуждения всегда имеют форму *приведения к абсурду*, но что сам этот приём, широко распространённый в математике, является не математическим, а адвокатским, обязательным в судах Древней Греции в эпоху, когда как раз складывалась геометрия. Можно думать, что он и в математику перекочевал из судоговорения.

Теперь конкретно о периодограмме, о выделении в ряду наблюдений $x_0, x_1, ..., x_n$ (или в непрерывной записи $x(t)$, $0 \le t \le T$) периодических особенностей. Для этого применяют тот или иной способ сравнения наблюдений с идеальной периодической функцией

$$e^{i\omega t} = \cos\omega t + i\sin\omega t$$

или с какой-либо другой периодической функцией. Набольшую сложность представляет определение скрытого периода или нескольких периодов в наших наблюдениях.

Для этого в математической статистике известен приём, состоящий в вычислении выражения

$$I_T(\omega) = \ | \int_0^T e^{-i\omega t} x(t) dt \, |^2 ,$$

или для дискретных наблюдений

$$I_T(\omega) = \ | \sum_{t=0}^{n} e^{-i\omega t} x_t \, |^2$$

при многих значениях ω и поиске одного или нескольких максимумов функции $I_T(\omega)$. На самом деле, см., например, книгу Тимофеева [19][2.16], в науке известно ещё несколько выражений, отличающихся от $I_T(\omega)$ пределами интегрирования (суммирования) и учётом не только модуля, но и аргумента комплексных величин. Получаются разнообразные графики, по поведению которых при разных ω можно локализовать возможные периоды, содержащиеся в наблюдениях. Строго говоря, рассмотрения Тимофеева не вероятностны: вероятностный подход требует привлечения понятия об ансамбле мыслимых реализаций наблюдений (из которого мы видим лишь одну реализацию) и оценки на основе этих представлений статистической значимости выделенных периодов. Это возможно сделать при определённых предположениях о модели, описывающей наблюдения. Например, очень удобно, если модель в дискретном случае имеет вид

$$x_t = \sum_{k=1}^{n} a_k \sin(\omega_k t + \varphi_k) + \xi_t , \qquad (2.3.1)$$

где $\xi_0, \xi_1, ..., \xi_n$ *независимые* одинаково распределённые случайные величины. В непрерывном случае подобная модель с независимостью значений *помехи* $\xi(t)$ при любых сколь угодно близких значениях t (*белый шум*) менее реалистична.

Можно привести ряд физических примеров, когда модель (2.3.1) реалистична. В качестве первого попавшегося примера годятся, скажем, наблюдения блеска переменной звезды, если он наблюдается не слишком часто. Разумность применения периодограммного метода и возможность определённых оценок значимости в таких случаях несомненна. Но в более сложных случаях, когда о вероятностной модели помех, искажающих процесс, говорить нельзя, вероятностно-статистические рассмотрения с оценками значимости невозможны.

Основываясь главным образом на применениях метода периодограммы к рядам экономических наблюдений, автор сделал ряд скептических замечаний по поводу реально достигнутых успехов [ii]. Эти замечания и вызвали то единственное бесспорное с научной точки зрения возражение, о котором говорилось здесь во Введении. Оно последовало от Тимофеева.

Он указал весьма интересный и неожиданный пример практического применения периодограммы. Повторяем, что в данном случае исследование не носит вероятностного характера (на периодограмме выделяются некоторые бесспорные пики, оценка значимости которых не нужна), но мы всё же кратко опишем его, следуя материалам Тимофеева.

Рассматриваемая проблема возникла при попытке автоматизации процесса непрерывного выщелачивания боксита под давлением. Установка состоит из ряда соединённых друг с другом особых сосудов, *автоклавов* больших размеров (диаметром 3 *м* и высотой 14 *м*). Через них последовательно проходит пар и пульпа. Расход пара сильно колеблется, тогда как температуру смеси желательно поддерживать постоянной. Обычный регулятор, предназначенный для этой цели, следовало *настроить*, но эта операция не удавалась. Поэтому диаграмму расхода пара было решено анализировать пероидограммно. Анализ показал наличие двух скрытых периодов, 33,8 *мин* при амплитуде 2,44 *т* пара в час и 6,5 мин при амплитуде 0,63 *т* пара в час. Принимая совпадение этих гармоник по фазе, можно рассчитать затухание системы, которым она должна обладать для погашения такого возмущения, и соответственно выбрать уставки регулятора.

Эти данные были сообщены наладчику, который и установил требуемые деления на регуляторе, после чего система вошла в режим практически постоянной температуры. Результат показался почти чудесным, и поэтому попытались проверить, не сможет ли система работать при более или менее произвольных уставках регулятора. Но эти попытки немедленно приводили к нарушению нормального режима.

Наладка регулятора сходного агрегата обычным путём проб потребовала его выключения в течение двух смен (ранее он работал при ручном управлении), что означало существенный убыток. Скрытые периоды на первом агрегате остались теми же самыми через два года, но при несколько изменённых амплитудах гармоник.

Итак, если имеется запись возмущающего воздействия на входе системы, то выделение из неё периодических составляющих может позволять выбор рациональных параметров регулятора. Поскольку рассматриваемая система чрезвычайно сложна, в частности, заведомо нелинейна, то её исчерпывающего анализа заранее провести нельзя, так что окончательный вывод об успехе метода мог быть сделан лишь на основе практического результата.

3. Заключение

Некоторые проблемы современного развития теории вероятностей

Примеры, приведённые в предыдущей главе, имеют целью иллюстрировать ту мысль, что вопрос о границах применимости теории вероятностей не может быть решён умозрительно с помощью логического обоснования или обоснования противного. Единичный практический успех тоже не даст научной уверенности в правоте теоретической концепции (вспомните успех австралийского дикаря). Только длительные исследования на протяжении многих лет (уже почти 20 лет ведётся Фремингемское исследование) и даже многих поколений учёных (как, к примеру, исследование наследственности, начатое Менделем) дают надёжный результат. В чисто методическом смысле такие исследования дают полную возможность экспериментальной проверки многих предположений вероятностно-статистического характера. В частности, могут быть проверены статистическая однородность (например, непараметрическими критериями различия двух эмпирических функций распределения), доверительные интервалы (напомним, что мы отвергли доверительный интервал для массы Юпитера, который принял Лаплас, см. конец § 1.1) и многое другое.

Итак, неверно, что этим предположениям не угрожает никакая экспериментальная проверка (это возражение Ю. И. Алимову). Впрочем, если бы просто собрать статистический или не статистический ансамбль всех случаев, когда применяются вероятностные методы, и подсчитать, в каком проценте случаев прав Алимов, а в каком – мы, то можно опасаться, что Алимов набрал бы подавляющее большинство голосов. Нам пришлось бы прикрыться тем доводом, что в науке количественное большинство голосов может и ничего не значить.

Но все эти рассуждения касаются одной стороны вопроса: что может теория дать практике и при каких условиях. Попытаемся теперь подумать, что может, обратно, практика дать теории. Для математики это очень старый вопрос, и по нему, разумеется, существуют самые крайние мнения. Для начала процитируем известного французского математика Ж. Дьедонне (1966, с. 11):

Напоследок я хотел бы подчеркнуть, сколь мало новейшая история оправдывает благочестивые пошлости прорицателей краха, регулярно предупреждающих нас о гибельных последствиях, которые математика неминуемо навлечёт на себя, если откажется от применений к другим наукам. Я не собираюсь утверждать, что тесный контакт с иными областями, такими, как теоретическая физика, не выгоден для обеих сторон. Но совершенно ясно, что из всех поразительных достижений, о которых я рассказывал, ни одно, за возможным исключением теории распределений, ни в малейшей степени не пригодно для физических применений. Даже в теории уравнений с частными производными упор сейчас гораздо больше на <u>*внутренние*</u> *и структурные проблемы, чем на вопросы, имеющие прямое физическое значение. Даже если бы математика насильно была бы отрезана от всех прочих каналов человеческой деятельности, в ней достало бы на столетия пищи для размышления над большими проблемами, которые мы должны ещё решить в нашей собственной науке*.

Что можно возразить? Прежде всего, раз речь идёт о времени порядка нескольких столетий, то не мешает обратиться к истории и посмотреть, нет ли примеров того, что происходит с теми или иными областями интеллектуальной деятельности, интерес к которым сохраняется в указанном интервале времени. Примером такой активности является средневековая схоластика.

Схоласты были умными, трудолюбивыми людьми, Во всяком случае, объём их сочинений по порядку такой же, как, скажем, сочинений Лапласа (количество бумаги, которую может исписать человек за свою жизнь, видимо, не очень зависит от содержания написанного). Университеты и академии были первоначально созданы для схоластов ввиду важности тех морально-этических приложений, которое имело, или делало вид, что имело, их творчество. Из этих учреждений никто их не *изгонял калёной метлой*, но как-то само собой получилось, что место схоластов заняли учёные, физики, математики, химики [алхимики?, астрономы] и пр. Почему же это произошло?

На наш взгляд, причина в том, что схоластика постепенно замкнулась в своих внутренних делах, перестав давать обществу что-либо в виде решения существенных для жизни нравственных вопросов. Например, как в средние века, так и теперь каждый решает для себя вступать или не вступать в брак. Схоласты, естественно, обсуждали этот вопрос, но их ответ превращался в длинную сводку мнений святых отцов и древних философов, одни из которых думали так, а другие иначе. Так что же делать *практически* работающему священнику, которому этот вопрос задаёт прихожанин? Постепенно эти вопросы стали считаться в серьёзной *науке* неприличными, а затем как-то само собой общество стало считать неприличной и самуЮ схоластику[3.1].

Этот пример заставляет с осторожностью думать о том, что случилось бы с математикой, будь она в течение столетий отрезана от всех прочих каналов деятельности. Механизм явления был бы очень прост: постоянно сокращалось бы число молодых людей, желающих посвятить себя математике, ибо как раз эти прочие каналы деятельности играют важнейшую роль в привлечении интересов молодёжи.

Но в конце концов (и статья Дьедонне убеждает нас в этом) можно допустить, что существует математика разных видов: одна направлена на свои собственные интересы, другая – на приложения. Оба вида имеют полное законное право на существование, потому, например, что, скажем, на основе теории функций действительной переменной (ТФДП, явно первого вида) возникла колмогоровская аксиоматика теории вероятностей, необходимая в определённом смысле и для приложений. Но в таком случае спрашивается: к какому виду математики относится теория вероятностей?

Отличительная черта математики первого вида – некоторое особое изящество при сравнительной простоте, которая не делает невозможным восприятие этого изящества[3.2]. В теории вероятностей есть довольно много результатов именно такого вида, большей частью связанных с колмогоровской аксиоматикой и напоминающих ТФДП. Но основное содержание этой науки, сложившееся во времена Лапласа, развивавшееся после него и разрабатываемое в настоящее время, увы! никак не прекрасно.

Например, предельные теоремы обычно формулируются довольно благопристойно, но, их доказательства как правило, безнадёжно длинны, трудны и запутанны. Их единственный смысл состоит в получении сравнительно простых вероятностных распределений в надежде на описание некоторых реальных явлений. Обращение к реальности всегда освежает, тогда как разрыв связей с ней опасен схоластическим перерождением.

Математика прекрасна, но её отдельные представители способны переродить её, например, в схоластику. Перейти грань, за которой начинается схоластика, к сожалению, в математике очень легко. Внутренние вопросы очень притягательны, всегда хочется навести порядок в своём собственном доме, потому что это свой дом и притом это всего доступней. Так где же находится та опасная грань, перейдя которую мы займёмся лишь натиркой полов в своей собственной квартире, но ничего не дадим обществу? Эта грань хорошо заметна лишь в исторической перспективе, но в каждый данный момент времени она крайне неопределённа и зыбка.

Удивительным образом здесь частично может помочь статистика, на этот раз в обличье наукометрии (Налимов и др., 1969). Сравнительно недавно был применён простой приём, состоящий в формальном исследовании цитирований, т. е. списков литературы, прилагаемых к каждой научной статье. Количество людей, интересующихся данным кругом вопросов, можно примерно оценить, поняв, какие группы авторов статей цитируют друг друга. Разумеется, если на этом попытаться построить какую-то политику административных оценок, все авторы начнут чисто формально цитировать друг друга, а разобраться, цитировано ли по существу, или лишь в обмен на такую же услугу, практически невозможно. Но пока нет никакого административного давления, исследование цитирований является ценным и более или менее объективным приёмом наукометрии[3.3]. И вот такой анализ показывает, что в теории вероятностей цитируют друг друга лишь весьма немногочисленные, по сравнению, скажем, с физикой[3.4], группы авторов. Это означает сужение интересов, в значительной мере обусловленное громоздкостью применяемого математического аппарата, что типично для схоластического перерождения.

Едва ли не простейший путь борьбы с такой опасностью состоит в обращении к физическим приложениям. Их серьёзность никогда и никем не оспаривалась. Ныне интересы сосредотачиваются, в частности, вокруг проблем статистической физики. Здесь применяется весьма широкий и сложный математический аппарат, способный удовлетворять разнообразные вкусы. Физические проблемы хороши ещё тем, что средствами математики удаётся очень многое сделать в них.

Но гораздо разнообразнее область, так сказать, чисто статистических приложений теории вероятностей. В очень многом и важном далеко идущий математический анализ невозможен. Но обширный эмпирический материал может в необходимой мере возместить скудость теоретических представлений. Во всех таких случаях статистическая обработка является одним из главных средств исследования. Мы здесь привели примеры именно такого рода, совершенно не касаясь бесспорного вопроса о физических приложениях теории вероятностей.

В чисто статистических вопросах основную роль играет та или иная вероятностная модель явления. В простейшем случае под моделью может пониматься просто предполагаемый вид закона распределения наблюдений (нормальный, показательный, вейсбулловский[3.5] и пр.) В более сложных случаях модели усложняются. Известны, например, теории надёжности и массового обслуживания, которые часто рассматривают довольно сложные аналитические модели.

Некая диспропорция современного развития теории вероятностей состоит в том, что теоретических моделей, даже достаточно исследованных аналитически, накоплено немало. Но во многих случаях их практические сопоставления с действительностью ни разу не делались. Конечно, создание теоретической модели есть необходимый начальный этап, без которого никакое сопоставление с действительностью и никакое понимание фактических данных вообще невозможно. Однако, слишком часто исследование останавливается на этом начальном этапе. Между тем, каждое сопоставление с действительностью обычно вызывает к жизни новые модели, т. е. и в этом смысле действует освежающе.

Библиография автора

1. Р. фон Мизес (1930), *Вероятность и статистика*. М. – Л.

2. А. Н. Колмогоров (1936), *Основные понятия теории вероятностей*. М. Второе издание.

3. Брошюра автора [i].

4. В. Н. Тутубалин (1972), *Теория вероятностей*. М.

5. Ю. И. Алимов (1974), Про застосування методи в математичной статистики до обрабки експериментальних данных. *Автоматика*, № 2, с. 21 – 33.

6. М. Планк (1966), *Единство физической картины мира*. Сборник статей. М. Ссылка неясна.

7. А. А. Боровков (1972), *Курс теории вероятностей*. М.

8. А. В. Гулыга (1975), Может ли наука быть безнравственной? *Природа*, № 12, с. 45 – 49.

9. С. Н. Давиденков (1975), Психофизиологические корни магии. *Природа*, № 8, с. 68 – 78.

10. А. Н. Колмогоров (1940), Об одном новом подтверждении законов Менделя. *Докл. АН СССР*, т. 27, № 1, с. 38 – 42.

11. Л. Н. Большев, Н. В. Смирнов (1967, 1968), *Таблицы математической статистики*. М.

12. Т. К. Енин (1939), Результаты анализа расщепления гибридов томата по отдельным семьям. *Докл. АН СССР*, т. 24, № 2, с. 176 – 178.

13. Брошюра автора [ii].

14. В. А. Тимофеев (1960), *Теория и практика анализа результатов наблюдений* и т. д. *Тр. ЛЭТИ*.

15. В. А. Тимофеев (1973), *Математические основы технической кибернетики.* (Конспект лекций.). Пенза.

16. В. А. Тимофеев (1975), *Инженерные методы расчёта и исследования динамических систем*. Л.

17. Ж. Дьедонне (1966), Современное развитие математики. *Математика* (сб. переводов), с. 3 – 11.
18. В. В. Налимов, З. М. Мульченко (1969), *Наукометрия.* М.
Ермолаева Н. И. (1939), Ещё раз о *гороховых законах*. *Яровизация*, № 2, с. 79 – 86. Автор уничижительно назвала законы Менделя.
Толстой Л. Н. (1910), *Путь жизни. Полн. собр. соч.*, т. 45.
Laplace P. S. (1812), *Théorie analytique des probabilités. Oeuvr. Compl*., t. 7. Paris, 1886.
Лаплас П. С. (1814, франц.), *Опыт философии теории вероятностей*. Русский перевод 1908 г. перепечатан: Прохоров Ю. В., редактор (1999), *Вероятность и математическая статистика. Энц*. М., с. 834 – 863.
Truett J., Cornfield J., Kannel W. (1967), A multivariate analysis of the risk of coronary heart disease in Framingham. *J. Chronic Diseases*, vol. 20, pp. 511 – 524.

Примечания

0.1. Это неверно, см. наш комментарий к брошюре [ii].

0.2. Проще было бы попросить комментировать подобные мнения.

0.3. Особенно скверно отражается эта традиция при переводах отечественной журнальной литературы. В достойных иностранных изданиях от авторов требуют чётких разъяснений, а по собственному опыту знаю, что нет никакой охоты улучшать перевод небрежно написанного текста.

1.1. Ср. Лаплас (1814/1999, с. 856 лев. столбец): *Лейбниц всё ещё руководствуется странной и развязной метафизикой*.

1.2. Непрестанные ссылки автора на Лапласа неуместны. До Лапласа теория вероятностей развивалась как отрасль чистой науки, Лаплас же разумно сместил её в русло прикладной математики (и добился фундаментальных успехов в естествознании). Но уровень абстракции его теории вероятностей был низок, и её пришлось создавать заново. С точки зрения популяризации этой теории Лапласа, к сожалению, затмили поверхностные и даже противоречивые, но захватывающие сочинения Кетле.

1.3. *Внешнее описание* было всегда необходимо, наукометрия же не ограничилась той скромной ролью, которую автор отвёл ей. Она руководит современной наукой **законом кувалды**, и можно пожалеть, что национальные академии смирились с ней. См. наши публикации (2015, § 5 Предисловия и 2018).

1.4. Такого человека вообще никогда не было. Боудитч, переводчик монографии Лапласа на английский язык, чётко заметил, что ему приходилось тратить *часы и может быть дни*, чтобы понять его некоторые рассуждения (Шейнин 2019, § 8.4).

1.5. Это необоснованное утверждение совершенно неверно.

1.6. Вот соответствующие ссылки на Лапласа (1814/1999): о Паскале, с. 849 лев. и 859 лев. О (Френсисе) Бэконе с. 860 прав. О Лейбнице с. 856 лев.

1.7. Фраза крайне неудачна.

1.8. Об ошибках Лапласа и его нежелании признать превосходство Гаусса в теории ошибок см. Шейнин (2019, § 8.4). Но свою непростительную ошибку он совершил в другой книге 1796 г. (и не исправил в её последнем прижизненном издании 1813 г.), см. там же. Он заявил, что эксцентриситеты планетных орбит были вызваны разностями температур и давлений в различных местах планет, тогда как Ньютон разъяснил эти эксцентриситеты скоростью обращения планет около Солнца.

1.9. Ошибок действительно не было, так ведь и содержания в рассуждениях тоже было немного.

1.10. Лаплас (1812/1886, с. 516 – 519), однако, оговорился: его оценка сохранится, если наблюдения будут обрабатывать тем же методом. Мы бы добавили: метод наблюдения также должен был бы оставаться прежним.

1.11. Уже Гиппарх и Птолемей должны были представлять себе существование систематических влияний, хотя бы действия вертикальной рефракции.

1.12. Следовало бы учитывать возможность ошибочного оправдания виновного. И вообще все подобные рассуждения были у Пуассона.

1.13. Мы сказал бы *логическая* критика.

1.14. У истоков стояли Якоб Бернулли, недооценённый Бейес и Муавр.
1.15. У Мизеса не было действительной аксиоматики (Хинчин 1961).
1.16. Это описание следовало связать с законом больших чисел.
1.17. Это даже невозможно.
1.18. Это непонятно.
1.19. У Мизеса по сути не было аксиоматики (Прим. 1.15), а потому выражение *вторая аксиоматика Колмогорова* непонятно. С другой стороны, автор упомянул теоретико-множественную аксиоматику, т. е., возможно, аксиоматику теории множеств, которая восходит к Цермело, 1906. О Жане Вилле (1910 – 1989) можно прочесть, например, в английской Википедии и в первой главе книги Shafer, Vovk (2001), и мы можем также назвать книгу Постников (1960), но о *концепции Вилле – Постникова* нам ничего не известно.
Наконец, можно по крайней мере проверять полноту и непротиворечивость предпосылок.
1.20. Что же в данном случае означает *методическая цель*?
1.21. См. по крайней мере конец Прим. 1.19.
1.22. См. Прим. 1.6.
1.23. Именно так полагал Лаплас (Шейнин § 8.2-1).
2.1. Фраза крайне неудачна, ср. Прим. 1.7.
2.2. Переходы от *сигмы* к *дельта* и обратно следовало разъяснить.
2.3. *Гистограмма* это не способ. Ср. Прим. 1.7 и 1.22.
2.4. Нет никаких пояснений.
2.5. Как было намечено в [ii, Прим. 2.3], этот анализ в более общем смысле называется предварительным исследованием данных.
2.6. Это неясно.
2.7. Ср. *Новый Завет*, Матфей 24:36 и Марк 13:32.
2.8. Грудная жаба – устаревшее название стенокардии.
2.9. Ряд значений (2.2.1) не имеет никакого отношения к факторам риска.
2.10. Обследование совсем не было *самым простейшим*. Достаточно обратить внимание на регистрацию кровяного давления, которая несомненно требовала большого числа измерений и их грамотной обработки.
2.11. Этот процент относился к мужчинам, у женщин он был существенно меньшим, см. выше.
2.12. Здесь и в других местах данные приведены с непомерным числом значащих цифр.
2.13. Буквой Ф обычно обозначается нормальное распределение. Читатель должен догадываться о её значении здесь.
2.14. Приведённые числа не могут ни о чём свидетельствовать.
2.15. *Графическое вычисление*: весьма необычное выражение
2.16. Ссылка неверна.
3.1. В схоластике мнение каждого отца церкви считалось вероятным, что привело к появлению *пробабилизма*. Его позднейшим аналогом стала неаддитивная вероятность у Якоба Бернулли. О колебаниях при вступлении в брак в эпоху заката схоластики см. эпопею женитьбы Панурга у Рабле.
3.2. О чистой и прикладной математике см. Грекова (1976). Она, кстати, процитировала подходящий научный фольклор:
Чистая математика делает то, что возможно, так, как нужно.
Прикладная математика делает то, что нужно, так, как можно.
3.3. См. Прим. 1.3.
3.4. Следовало бы сравнивать не со всей физикой, а с каким-либо её разделом.
3.5. Распределение Вейсбулла характеризует многие явления, например, выживаемость, надёжность, перемены погоды.

Библиография публикатора

И. Грекова (1976), Методологические особенности прикладной математики и т. д. *Вопр. философии*, № 6, с. 104 – 114.
А. Г. Постников (1960), *Арифметическое моделирование случайных процессов*. М.
А. Я. Хинчин (1961), Частотная теория Мизеса и современные идеи теории вероятностей. *Вопр. философии*, № 1, с. 91 – 102; № 2, с. 77 – 89.
О. Б. Шейнин, O. Sheynin (2015), *Чёрная книга*. Берлин. **S, G,** 69.

--- (2018), History of mathematics: some thoughts about the general situation. *Silesian Stat. Rev.*, No. 16 (22), pp. 127 – 156.

--- (2019), *Теория вероятностей. Исторический очерк*. Берлин. **S, G,** 11.

G. Shafer, V. Vovk (2001), *Probability and Finance. It's Only a Game.* New York.

IV

В. Н. Тутубалин

Ответ на критические замечания Ю. И. Алимова в связи с проблемами приложения теории вероятностей

Автоматика, № 5, 1978, с. 88 – 90

В № 1 за этот год опубликована статья Алимова [1], в которой критически рассматриваются некоторые мои публикации. Эта статья является единственным известным мне откликом на мои работы учебно-методического и научно-популярного содержания. Поскольку дискуссии, в том числе печатные, являются необходимейшим условием развития науки и преподавания, следует весьма положительно оценить инициативу этого журнала и значительный (как будет, в частности, видно из дальнейшего) труд Алимова, благодаря которым только и стала возможной эта дискуссия.

Алимов хорошо известен ввиду ряда своих публикаций, главным образом критического свойства по проблемам применений теории вероятностей. Думаю, что общие задачи его публикаций мало отличаются от тех, которые я ставлю себе: мы оба, видимо, согласны в том, что количество лжи, которое происходит с помощью *применений* теории вероятностей, так велико, что даже и не может быть терпимо. В исторической перспективе одно только выражение этого факта в печати является вполне действенным средством борьбы с указанным злом. Да, собственно, в обществе сейчас происходят глубинные процессы, благодаря которым резко возрастает роль моральных начал. Этому-то мы с Алимовым и обязаны известной популярностью наших публикаций – в той, не слишком большой мере, в какой мы её имеем. В противном случае вопрос о популярности публикаций вовсе не стоял бы: их бы просто не было.

Таким образом, связывая вопрос об истинности научных работ прежде всего с уровнем общественной морали, я довольно скептически смотрю на возможность решения этой проблемы чисто научными средствами, скажем, введением изложения теории вероятностей по Мизесу вместо общепринятой сейчас аксиоматики Колмогорова. Я уж не упоминаю идею административных мер порицания, которую высказал Алимов [1, с. 82]. Эта идея была бы реальной лишь, если лица, облекаемые административной властью, автоматически облекались и знанием истины или хотя бы стремлением к ней.

Кстати, мне хотелось бы обратить внимание Алимова на то обстоятельство, которое я ранее заметил на собственном опыте. Любая попытка переложения или цитирования взглядов других лиц, видимо, вносит неизбежные искажения[1]. Например [1, с. 74]:

Сравнивая подход Мизеса с мёртвым языком, Тутубалин тем не менее отмечает …

А на самом деле у меня сказано [2, с. 148]:

В целом, теперешнее отношение специалистов к языку теории Мизеса можно сравнить с отношением к мёртвому языку, на котором почему-то никто говорить не хочет, но при соответствующих поправках и переделках, вполне можно было бы сказать всё то, что говорят на живом языке.

Таким образом, моё дружелюбное отношение к теории вероятностей по Мизесу в результате переложения (вполне корректного, но только краткого) совершенно исчезло и заменилось пренебрежением.

Этот пример вместе с моим общим мнением о том, что искажения такого рода практически неизбежны, достаточно объясняют, почему я подробно отвечаю на каждый пункт критики Алимова. В том, что касается общих высказываний, дело сводится к выбору того или иного оттенка мысли. Например, если Алимов считает, что закон больших чисел является предельной теоремой, и решает скромную задачу теории вероятностей, не связанную непосредственно с принципиальными вопросами возможности её приложения, то пусть он будет прав.

Гораздо интереснее обратиться к конкретным примерам применений теории вероятностей, ибо они всегда богаче общих рассуждений. Например, крупные физики, создававшие эту науку, обычно сами осмысливали её философски и не нуждались в услугах философов. Я не отрицаю общественно полезной роли философов, но их потребитель – не ведущие учёные, а массы тех, кто не занимает или ещё не занимает ведущей роли в науке. Основная заслуга Алимова как критика состоит, как мне кажется, в том, что он рассмотрел конкретный цифровой материал.

Речь идёт об экспериментальной проверке простейшего менделевского закона распределения признаков в отношении 3:1. Экспериментальный материал дан в двух статьях Ермолаевой [3] и Енина [4], причём первый автор являлся представителем школы Лысенко, а второй, её противником. Колмогоров [5] опубликовал подробный анализ результатов Ермолаевой с тем выводом, что вместо опровержения менделевского закона она полностью его подтвердила. Там же Колмогоров дал понять, что результаты Енина сомнительны, потому что они слишком хорошо подтверждают закон Менделя, но не привёл их подробного анализа. В научно-популярной брошюре [6] я счёл целесообразным напомнить об указанной работе Колмогорова и дополнить её собственной обработкой результатов Енина.

Алимов, во-первых, обработал те же данные иным способом, а во-вторых, сделал ряд возражений. Они почему-то были адресованы только мне, хотя часть из них в равной степени затрагивает обработку, сделанную Колмогоровым. Я начну с того возражения, которое я понял и считаю серьёзным, затем коснусь тех возражений, которых я либо не понял, либо не считаю серьёзными.

Алимов замечает, что *семьи*, рассмотренные Ермолаевой, во многих случаях невелики (не более десятка наблюдений). В таком случае нормальное приближение для частот проявления определённого фенотипа, рассмотренное Колмогоровым, должно быть очень грубым. В частности, наличие нормированных частот, меньших – 3, которое я [6] оценил как значимое отклонение от закона Менделя, может, согласно Алимову, объясняться асимметрией биномиального закона. Алимов объявляет мой вывод неверным (это несколько торопливо, следовало бы объявить его лишь необоснованным) и объясняет мою ошибку, которую в состоянии избежать *каждый студент втуза*, *общей аберрацией понятий, возникающей при использовании немизевского языка и ритуалов математической статистики*.

Но на самом деле всё гораздо проще. При подготовке брошюры [6] я не знакомился с работой Ермолаевой из-за некоторых неудобств, связанных с получением нужного журнала. Теперь, когда эти данные стали предметом дискуссии, я посмотрел этот источник. Данные по расщеплению в отдельных семействах представлены в ней в табл. 4 и 6. В табл. 4 семейства пронумерованы номерами о 1 до 100, но по неизвестной причине пропущены номера 50 и 87. В табл. 6 нумерация начинается с 22 и идёт до 148, но пропущены номера 92, 95, 115, 127 и 144. Однако, в итоговой таблице в первом случае считается 100 семейств, а во втором – 127.

В своей работе Колмогоров ехидно заметил, что в первой таблице он насчитал 98 семейств, а во второй 123. Он сам ошибся на единицу, так как во второй таблице всего 148 – 21 – 5 = 122 семейства. Общий стиль работы Ермолаевой, прямо скажем, гнусный. Непонимание смысла вычисляемых *биометрическими методами* ошибок для численности расщеплений, очевидно. Совершенно понятно, что её данные не заслуживают особенно серьёзного отношения. Но если говорить лишь об анализе таблиц Ермолаевой в том виде, в каком они есть, то всё-таки Алимов неправ, упрекая меня в том, что я обнаружил отклонение от законов Менделя там, где его нет. Дело в том, что в табл. 4 имеется результат расщепления 0:17, а в табл. 6, 0:10 (вместо ожидаемого 3:1), вероятности которых равны 4^{-17} и 4^{-10}, так что при 200 с небольшим испытаниях такие события не могли произойти.

Хотелось бы заметить, как по поводу обработки у Колмогорова, так и по поводу моей, что, возможно вопреки мнению Алимова, правильные научные результаты часто (?) получаются не за счёт того, что всё делается правильно, а ввиду особого везения.

Смысла возражений Алимова по поводу вычисления уровня значимости я не понял. Исстари, т. е. со времён Лапласа, пытаются, получив некое расхождение с теорией, вычислить, если возможно, как велика вероятность такого же или ещё большего расхождения, если теория правильна[2]. Если эта вероятность велика, скажем, равна 1/2, то всё в порядке. Если же мала, скажем, порядка 1/1000, то целесообразно поискать причину расхождения. Если она умеренна, скажем, порядка 1/10, то возникает сомнительный случай, когда принять окончательное решение невозможно. Можно ли возражать против такого использования уровня значимости?

Неясно мне также понимание независимости по Алимову. На с. 80 даётся понять, что *вторичные* испытания, т. е. данные о расщеплении признаков в разных семьях, могли бы быть статистически зависимы. Но как могут быть статистически зависимы исходы разных, никак не связанных друг с другом опытов? Я понимаю, что если в результате *одного* опыта могут наступить или не наступить события А и В, то они могут быть статистически зависимы, и трактовать эту зависимость по Мизесу надо с помощью одного протокола. Но если есть два совершенно разных опыта, то надо, видимо, привлечь нечто вроде прямого произведения двух протоколов.

Наконец, по поводу моей обработки данных Енина. Алимов замечает, прежде всего, что число семей у Енина (11 + 14 = 25) так невелико, что на обработку этих данных не стоило тратить ни времени, ни бумаги со специальной нелинейной шкалой. На это отвечу, что, напротив, я ставил своей целью показать, что изображение эмпирической функции распределения, в отличие от изображения гистограммы, позволяет получить осмысленные результаты даже при столь небольшом объёме выборок.

Далее, Алимов утверждает, что те выводы, к которым я пришёл, можно было бы получить путём составления удлиняющейся выборки. Это в известной мере и так, и не так. Для выборок примерно одинакового объёма по сравнению с аналогичными частотами Ермолаевой частоты во второй выборке Енина ближе к теоретическим. Это видно из таблицы [1, с. 78]. Если данные Ермолаевой принять за эталон, то может быть сделан вывод о неблагополучии с данными Енина. Но если вычислить статистику хи-квадрат [6], то нет нужды в эталоне.

По существу, на с. 80 – 81 Алимов выражает то мнение, что данные Ермолаевой должны были бы обрабатываться не с помощью модели нормального распределения для нормированных частот, а с помощью более тонкой статистической модели. В принципе я совершенно согласен, но только в качестве модели надо брать не смесь биномиальных распределений [1, с. 80, формула (21)], а прямо рассмотреть фактические численности семей. Получится серия биномиальных испытаний с известными численностями последних и известной вероятностью успеха. Аналитически эта модель мало удобна, так что глупейший метод Монте Карло[3], пожалуй, окажется наиболее эффективным для вычисления разных вероятностей, связанных с такой моделью. Так может быть найдено, например, истинное распределение статистики Колмогорова или какой-либо другой статистики, измеряющей отклонение от законов Менделя. Поскольку такие статистики довольно разнообразны, то мы приходим к выводу, что с научной точки зрения неисчерпаем не только электрон или атом, но неисчерпаемы и небрежно составленные таблицы Ермолаевой.

Библиография

1. Алимов Ю. И. (1978), О проблемах приложения теории вероятностей, рассмотренных в работах В. Н. Тутубалина. *Автоматика*, № 1, с. 71 – 82.

2. Тутубалин В. Н. (1972), *Теория вероятностей*. М., МГУ.

3. Ермолаева Н. И. (1939), Ещё раз о *гороховых законах*. *Яровизация*, № 2 (23).

4. Енин Т. К. (1939), Результаты анализа расщепления гибридов томата по отдельным семьям. *Докл. АН СССР*, № 2, т. 24, с. 176 – 178.

5. Колмогоров А. Н. (1940), Об одном новом подтверждении законов Менделя. *Докл. АН СССР*, № 1, т. 27, с. 38 – 42.

6. Брошюра [iii].

Примечания

1. Это слишком сильное утверждение. Впрочем, сам автор даже не заметил высказываний Пуассона о судебной статистике и в своих брошюрах неоднократно и совершенно неверно оценивал соответствующие результаты великого учёного.

2. Первые оценки расхождений теории и опыта появились у Гершеля, если не раньше, т. е. до Лапласа. Впоследствии, совершенно независимо от Лапласа, их неоднократно применяли в теории ошибок.

3. Странное и совершенно неверное мнение.

Публикатор: Основные идеи Ю. И. Алимова [1]

С. 71. Устойчивость исходных средних не доказана и должна быть обоснована экспериментально.

С. 73. Метод наибольшего правдоподобия (МНП) иногда всё ещё считается мостом, соединяющим теорию вероятностей с практикой. В контексте (с. 72) автор отрицает это утверждение, поскольку статистическая устойчивость испытаний должна быть доказана.

С. 73. Близость эмпирической частоты и начальной вероятности должна измеряться.

С. 74. Практика не следует за Мизесом, как полагает В. Н. Т., скорее произошло обратное.

С. 75. В статистике значимость МНП и других предельных (?) теорем сводится к решению обычной проблемы.

С. 76. Пояснение независимости испытаний не является фундаментально важным для подхода Мизеса. Статистическая независимость может быть выявлена самыми различными, в том числе периодическими последовательностями испытаний.

С. 77. Для приложений, переход от эмпирической частоты к вероятности является незаслуженной тратой строгой формализации, но не существенной чертой подхода Мизеса.

С.77. Без должного обоснования, но в соответствии с прежним утверждением, автор заявляет, что *так называемые усиленные законы больших чисел* очень далеки от теории вероятностей.