С. Д. Пуассон

Исследования о вероятности приговоров в уголовных и гражданских делах

Перевод О. Б. Шейнина

S.-D. Poisson
Recherches sur la probabilité des jugements
en matière criminelle et en matière civile.
Paris, 1837

С. Д. Пуассон
Исследования
о вероятности приговоров
в уголовных и гражданских делах

Перевод О. Б. Шейнина

Berlin, 2013

Текст настоящей книги размещен в Интернете
www.sheynin.de

ISBN 978-3-942944-29-8

Обложка и макет: Вячеслав Демидов (Viatcheslav Demidov)

NG Verlag (Viatcheslav Demidov Inhaber)
Berlin
Tel.: 030/4442460
Fax: 030/44739165
E-mail: SlavaDemidov@t-online.de
Internet: www.ng-verlag.de

Содержание

От переводчика ....
Исследования о вероятности приговоров в уголовных и гражданских делах, 1837...
[Аннотированное] оглавление...
Предисловие...
Гл. 1. Общие правила о вероятностях................................
Гл. 2. Общие правила, продолжение. Вероятности причин и будущих событий, выведенные из наблюдений прошедших событий ....
Гл. 3. Исчисление вероятностей, зависящих от очень больших чисел..
Гл. 4. Продолжение исчисления вероятностей, зависящих от очень больших чисел ....
Гл. 5. Приложение общих правил вероятностей к решениям судов присяжных и суждениям трибуналов........

От переводчика

Формально это издание является вторым, фактически же первым. По вине типографии в первом издании оказалось много слишком поздно замеченных ошибок, и все его экземпляры уничтожены, остался от него только номер ISBN.

Мы перевели основное теоретико-вероятностное сочинение Пуассона, но выпустили многие преобразования, не имеющие прямого отношения к теории вероятностей и насыщенные формулами, которые нелегко переписывать заново. Мы также не проверили подчас недостаточно пояснённые выводы формул.

Пуассон (Предисловие, § 13) указал цель своих первых глав, посвящённых собственно теории вероятностей:

В четырёх первых главах содержатся правила и общие формулы исчисления вероятностей, что освобождает читателя от их поисков в других источниках и позволяет рассматривать некоторые другие вопросы […].

Теоретико-вероятностное наследие Пуассона описано во многих источниках, в том числе и нами (Шейнин 1978; Гнеденко и Шейнин 1978), см. также Bru (1981). В нынешнем переводе мы комментировали отдельные вопросы, также касающиеся теории вероятностей у Пуассона; вот перечень важнейших из них:

1) Случайная переменная. В § 53 Пуассон определил дискретную случайную переменную, которая появилась у него ещё раньше (1829, § 2.5). Подобные переменные были известны, их применяли Симпсон, Лаплас и Гаусс и даже ещё раньше, Пуассон же ввёл их формально, хотя и назвал их явно временным термином. Впрочем, он неизменно рассматривал урновые схемы.

2) Шансы и вероятности. Пуассон вводил и шансы, и вероятности, и мы не уверены, что он всегда последовательно различал эти понятия. Во всяком случае, позднейшие авторы начали применять гораздо более подходящие термины, объективные и субъективные вероятности. Во многих случаях Пуассон рассматривал субъективные вероятности, и иногда его выводы были малополезны, а один из его примеров в § 11 привёл к вероятности, равной 1/2, что означало, как он сам заявил в § 4, полнейшее недоумение (что, кстати, указывает и теория информации). То же можно сказать о знаменитой задаче Бертрана о длине случайной хорды.

3) Прямой и обратный закон больших чисел, центральная предельная теорема. Закон больших чисел у Пуассона оказался принципом, притом заведомо слишком широким, и многие десятилетия статистики понимали его в таком же смысле.

Муавр ввёл простейший вариант центральной предельной теоремы, которую затем усовершенствовал Лаплас. Многие авторы, в том числе и Лаплас, предложили нестрогие доказательства нескольких вариантов этой теоремы, Пуассон к тому же доказывал её методически нестрого (и основывал на нём вывод закона больших чисел в надлежащем ограниченном смысле).

Этот закон, понимаемый расширенном смысле, применялся для решения обратной задачи, в которой теоретическая вероятность не была известна (или не существовала), и её приходилось оценивать по наблюдениям или даже вводить. Обратная задача менее точна, чем прямая, но понял это только Бейес, а не Якоб Бернулли или Муавр (Шейнин 2011), Пуассон же во всяком случае не изучил этого количественно.

4) Причины изучаемых событий. Пуассон неизменно вводил и их, и вероятности и силу их действия. Ни математики, ни статистики не продолжили этих изысканий; Кетле, например, изучал причины событий совершенно независимо от своего предшественника.

5) Математическая статистика. Изучая значимость эмпирических расхождений, т. е. занимаясь ещё не существовавшей математической статистикой, Пуассон фактически применял состоятельные оценки. Он также ввёл нулевую гипотезу и в качестве основного предположения, и как критерий случайности указанных расхождений. Его бывший студент, Gavarret (1840), заявил, что эти гипотезы, понимаемые во втором смысле, необходимы. И он, и Пуассон ограничивали приложение статистики (Пуассон ни разу не применил этого термина!) наличием большого числа наблюдений (Sheynin 2012).

6) Теория ошибок. Ввиду злополучного приоритетного спора Лежандра с Гауссом французские математики, включая Пуассона, обходили молчанием гауссову теорию ошибок. Умалчивал о ней и Лаплас, поскольку он разработал свой собственный вариант теории ошибок, – вряд ли применимой, поскольку она основывалась на большом числе наблюдений, а принятая им мера точности требовала вычислений, возможных только при нормальном распределении.

В нескольких местах Пуассон обсуждал угловые геодезические измерения, но его рассуждения оказались бесполезными ещё и потому, что он не был знаком с ними, не знал существенных обстоятельств полевых работ.

7) Три нововведения. Пуассон (§ 103) и ещё раньше (1811/1833, с. 637) ввёл функции, ныне называемые функциями Дирака.

Он (1829, § 1; 1837, с. 63 и 80) ввёл и функцию распределения и определил плотность как производную от неё. В XIX веке это нововведение применил А. Ю. Давидов в литографированном курсе своих лекций 1884 – 1885 гг. (Ондар 1971), а затем, как известно, А. М. Ляпунов.

Для полноты изложения добавим, что Пуассон (1824, с. 278) ввёл распределение, которое стало называться по имени Коши, и фактически выяснил, что оно устойчиво.

В целом, однако, теория вероятностей оставалась у Пуассона , как и у Лапласа, прикладной математической дисциплиной, а его сочинение, вопреки его принятому заглавию в немецком переводе (см. Библиографию), никак нельзя назвать учебником.

Пятая глава. Пуассон сам (Предисловие, § 8) так косвенно описал цель своей последней главы:

[…] *чёткая цель теории* – *вычисление для суда присяжных* […], *долей оправданий и осуждений, которые весьма вероятно будут иметь место, и шанса ошибки судебного решения.*

Пуассон ввёл вероятность ошибки присяжных и судей и, в отличие от Лапласа, предварительную вероятность вины подсудимого, а понятия *вины* и *невиновности* он решительно заменил на *подлежащие осуждению, оправданию*. Пуассон исследовал воздействие изменений в уголовном законодательстве на долю осуждений и на ошибки при осуждениях и оправданиях. Подобные труды относятся к предыстории понятий о критической области и ошибок первого и второго рода.

Пуассон извлёк статистические данные из нескольких источников, тем самым проделав большую работу. Он ни разу не упомянул ни зависимости между решениями отдельных присяжных (судей), ни неизбежных изменений во времени их отношения к делам одного и того же вида. И всё-таки многие его заключения были, видимо, более или менее верными.

Вопреки мнению Лапласа, приложения теории вероятностей к *моральным* наукам сильно критиковалось. Пуансо (Poisson 1836, с. 380), назвал применение исчисления вероятностей к *моральным вещам опасной иллюзией и ложным приложением математических наук.* Милль (1843/1914, с. 490) заявил, что

Неудачные приложения исчисления вероятностей […] *сделали* [его] *настоящим позором математики. Достаточно упомянуть о его приложении к установлению достоверности свидетелей и правильности приговоров, выносимых присяжными*.

В 1899 г., в связи с пресловутым делом Дрейфуса, Пуанкаре (Шейнин 1991, с. 167) положительно отозвался об этом утверждении, а несколько ранее (1896/1999, с. 22) заявил, что *наши привычки панургова стада* [баранов] противодействуют независимости суждений. Забыв, правда, о Кетле (и Курно), Heyde & Seneta (1977, с. 28 – 34) заметили, что в XIX в. в интересующей нас области *произошёл всплеск деятельности, стимулированный Пуассоном* (с. 31), но не обосновали этого утверждения. Они отметили возросшее ныне понимание важности общих данных о преступлении (например, о круге возможных преступников). Впрочем, многие учёные, включая Милля и даже Лейбница (в письмах Якобу Бернулли в самом начале XVIII в.), издавна придерживались того же мнения.

Курно (1843) включил в свою книгу главу об уголовной статистике и, как и раньше (1838), пытался исследовать влияние зависимости между присяжными или судьями. Несколько раз он (§§ 61, 93, 149 прим., 225 и 237) критиковал Пуассона, иногда косвенно, но сильно.

Неприятные обстоятельства

1) Книга плохо напечатана, и особенно это относится к формулам.

2) В ней много опечаток/ошибок, не замеченных автором. Мы указали по меньшей мере часть из них.

3) Формулы в первых двух главах не пронумерованы, в двух последующих главах нумерация явно недостаточна, но в последней главе она намного полнее. Мы пронумеровали много других формул (и применили более удобную систему нумерации).

Ссылки Пуассона на предыдущие формулы по необходимости затрудняют читателя, а в нескольких случаях они плохо понятны.

4) Предисловие не было разбито на параграфы, и мы сами проделали эту работу по своему разумению. Но Пуассон подразделил свой основной текст на сравнительно небольшие параграфы, что существенно облегчило чтение.

5) Пуассон явно предназначал свой труд для широкого круга читателей: некоторые пояснения слишком подробны, а выводы поясняются возможными соотношениями ставок пари. Но многие преобразования плохо понятны. Обсуждение петербургской игры (§ 25) малоудовлетворительно, а принцип Бейеса поверхностно описан в нескольких словах (§ 1 Предисловия).

6) Для современного читателя непременные выражения *величина A* и *дробь p* излишни (Пуассон всегда подразумевал правильную дробь), но мы не стали изменять их. Его терминология устарела, но понятна. Некоторые позднейшие обозначения, например $n!$, C_m^n, $\overline{p}$, естественно, отсутствуют.

В 1841 г. книга Пуассона вышла в немецком переводе с четырьмя, видимо, совершенно забытыми приложениями переводчика общим объёмом примерно в 150 страниц (пожизненные ренты и страхование жизни; моральное ожидание; вероятности средних результатов, т. е. Poisson S. D. (1824 – 1829); приложение теории вероятностей к натуральной философии). Чёткая печать перевода позволила нам проверить (разумеется, лишь в какой-то мере) правильность чтения многих формул Пуассона. Кроме того, мы сверили наш перевод некоторых фраз автора с соответствующим немецким текстом.

Сочинение Пуассона было дважды перепечатано (Париж, 2003, 2012).

Библиография

Гнеденко Б. В., Шейнин О. Б. (1978), Теория вероятностей. Глава в книге *Математика XIX века* [т. 1]. М. Ред., А. Н. Колмогоров, А. П. Юшкевич.

Ондар Х. О. (1971), О работах А. Ю. Давидова по теории вероятностей и его методологических взглядах. *История и методология естеств. наук*, вып. 11, с. 98 – 109.

Милль Дж. С. (1843, англ.), *Система логики*. СПБ, 1914. [М., 2011.]

Пуанкаре А. (1896, франц.), *Теория вероятностей*. Ижевск, 1999.

Шейнин О. Б., Sheynin O. B. (1978), Poisson's work in probability. *Arch. Hist. Ex. Sci.*, vol. 18, pp. 245 – 300.

--- (1991), Poincaré's work in probability. *Arch. Hist. Ex. Sci.*, vol. 42, pp. 137 – 172.

--- (2002), Sampling without replacement: history and applications. *NTM, Intern. Z. F. Geschichte u. Ethik Naturwiss., Technik, Med.*, Bd. 10, pp. 181 – 187.

--- (2011), Обратный закон больших чисел. *Историко-математич. исследования*, вып. 14 (49), с. 212 – 219.

--- (2012), Poisson and statistics. *Math. Scientist*, vol. 37, pp. 149 – 150.

Bru B. (1981), Poisson, le calcul des probabilités et l'instruction publique. In Métivier M. et al, Editors, *Poisson et la science de son temps*. Paris, pp. 51 – 94.

Cournot A. A. (1838), Sur l'applications du calcul des chances à la statistique judiciare. *J. Math. Pures et Appl.*, sér. 1, t. 3, pp. 257 – 334.
---- (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984. Editor, B. Bru.
Gavarret J. (1840), *Principes généraux de statistique médicale*. Paris.
Heyde C. C., Seneta E. (1977), *I. J. Bienaymé.* New York.
Poisson S. D. (1824), Sur la probabilité des résultats moyens d'observations. *Conn. des temps* pour 1827, pp. 273 – 302.
--- (1829), Sur la probabilité …Ibidem, pour 1832, pp. 3 – 22.
--- (1836, April 11 and 18), Note sur la loi des grandes nombres. *C. r. Acad. Sci. Paris*, t. 2, pp. 377 – 382, 395 – 400.
--- (1837), Sur la probabilité du tir a la cible. *Mémorial de l'artill.*, No. 4, pp. 59 – 94.
--- (1841), *Lehrbuch der Wahrscheinlichkeitsrechnung und deren wichtigsten Anwendungen*. Braunschweig. Übers. C. H. Schnuse.

Исследования о вероятности приговоров в уголовных и гражданских делах, 1837

[Аннотированное] оглавление

[Сводка] Исчисление вероятностей равным образом применимо к вещам любого рода, моральным и физическим и нисколько не зависит от их природы, лишь бы в каждом случае наблюдения предоставляли необходимые числовые данные.

Изложение *общего закона больших чисел*. Его подтверждение на многочисленных и разнообразных примерах, взятых из физических и моральных миров и позволяющих считать его в настоящее время опытным фактом, который никогда не будет опровергнут. Этот закон будет непосредственно доказан ниже.

Сводка данных наблюдений и результатов, которые получены в последней главе и относятся к вероятностям судебных решений по уголовным и гражданским делам.

Глава 1-я. Общие правила о вероятностях. Определение вероятности события. Различие, которое может быть сделано между *шансом* и *вероятностью*. Мера вероятности. Цель исчисления вероятностей. Доказательство главных правил этого исчисления. Примеры их применения[1] (§§ 1 – 13)

Формулы, относящиеся к повторению событий в серии испытаний. Решение задачи о *разделе ставки*. Решение другой задачи, основанное на разложении многочлена, возведенного в данную степень. Замечание о случае переменных шансов при испытаниях. Вероятность извлечь m белых и n черных шаров в выборке $(m + n)$ шаров без возвращения из урны, содержащей шары этих цветов в заданном соотношении (§§ 14 – 19)

Общее правило для определения вероятности составных событий, если шансы простых событий как-то меняются во время испытаний (§ 20)

Приложение исчисления вероятностей к определению выгоды от появления случайных вещей[2]. Вычисление различных шансов в прежней королевской лотерее Франции. Предрассудки противоположного смысла, касающиеся выхода номеров, притом равным образом плохо обоснованные, у игроков. Что понимать под *математическим* и *моральным ожиданиями*. Объяснение трудности в правиле математического ожидания[3] (§§ 21 – 25)

Существование неизвестного шанса, благоприятного для одного, неизвестно какого именно, из двух противоположных событий, всегда повышает вероятность повторения событий при двух или большем числе испытаний (§ 26)

Гл. 2-я. Общие правила, продолжение. Вероятности причин и будущих событий, выведенные из наблюдения прошедших событий. Смысл, который приписывается в исчислении вероятностей словам *причина* и *случай*. Правило для определения вероятностей различных возможных причин наблюденного события. Замечание о применении этого правила к последовательным событиям. Правило для определения по наблюденным событиям вероятностей других событий, которые зависят от тех же причин, что, однако, не предполагает никакого влияния появления прошлых событий на наступление будущих. Приложение этих двух правил к частному примеру (§§ 27 – 33)

Распространение тех же правил на случай, когда о событиях имеются некоторые предварительные сведения, полученные из наблюдений. Пример, подходящий для указания на необходимость обращать на это внимание (§§ 34 – 35)

Формулы, относящиеся к вероятностям свидетельских показаний. Случай, при котором требуется только знать истинно или ложно некоторое событие, удостоверенное или отрицаемое одним или многими свидетелями. Случай, при котором могут иметь место более двух событий и одно определенное из них удостоверено свидетелем. Теорема о вероятности события, о котором мы узнаем по традиционной цепочке свидетелей (§§ 36 – 40)

Когда возможно очень большое число событий, априорно имеющие равные и очень низкие вероятности, появление одного из тех, которые представляют что-либо *примечательное*, следует весьма вероятно приписать особой причине C, аналогичной, к примеру, человеческой воле, а не случаю. Если примечательные события перед наблюдением были намного более вероятны чем остальные, вероятность вмешательства причины C намного убывает и может быть настолько, что нет смысла обращать на нее внимание (§§ 41 – 42)

Преобразование формул, относящихся к вероятностям причин будущих событий, в определенные интегралы, когда число возможных причин бесконечно. Можно не рассматривать общие причины прошлых и будущих событий и считать те и другие составными событиями, зависящими от одного и того же простого события G, неизвестный шанс которого может принимать бесконечное множество значений (§§ 43 – 45)

Приложение этих интегралов к задаче, в которой событие G появляется m раз в $(m + n)$ испытаниях, а противоположное событие H, – остальные n раз и требуется определить вероятность, что они появятся m_1 и n_1 раз соответственно в $(m_1 + n_1)$ будущих испытаниях. Случай, когда априорно задано, что неизвестный шанс G очень мало отличается от заданной дроби[4] (§§ 46 – 48)

Изложение теоремы Якоба Бернулли о повторении событий при очень большом числе испытаний пропорционально их соответствующим шансам, известным или нет, но предположенным постоянными. Приложение к примеру, извлеченному из *Моральной арифметики* Бюффона. Указание на доказательство этой теоремы, основанное на формуле бинома (§§ 49 – 51)

Изложение трех общих положений, доказанных в гл. 4-й и относящихся к повторению событий, чьи шансы изменяются любым образом в течение испытаний. Вывод общего закона больших чисел, уже подтвержденного в Предисловии. Он выражен в двух уравнениях, которые являются основанием всех важных применений исчисления вероятностей (§§ 52 – 54)

Приложение первого уравнения к примерам. Существенное различие между применениями выведенных из наблюдений постоянного шанса и среднего шанса событий. Постоянная пропорциональность мужских и женских рождений. Отношения, которые должны существовать между совпадением и несовпадением полов первых новорожденных в одной и той же женитьбе (§§ 55 – 59)

Указание, в качестве приложения второго уравнения, на вычисление средних ошибок наблюдений, среднего срока жизни для различных возрастов и влияния ветров на высоту [уровня] моря (§§ 60 – 62)

Отступление о принципе *причинности*. Опровержение мнения Юма о простом *совпадении причин* и *действия*. Показано, что существование причины, способной *наверняка* вызвать явление, может иметь очень высокую вероятность, хотя бы оно и наблюдалось лишь малое число раз (§§ 63 – 64)

Вероятность существования и отсутствия постоянной причины некоторых явлений, соединенной с переменными причинами и со случаем и не всегда приводящей к их наступлению. Что следует понимать под *счастьем* и *несчастьем* в игре (§ 65)

Гл. 3-я. Исчисление вероятностей, зависящих от очень больших чисел. Необходимость прибегать к приближенным методам для вычисления значений произведения очень большого числа неравных сомножителей. Метод Лапласа для сведения в сходящиеся ряды функций очень больших чисел, вначале выраженных определенными интегралами. Приложение этого метода к произведению натуральных чисел $1 \cdot 2 \cdot 3 \cdot ... \cdot n$. Формула Валлиса (§§ 66 – 68)

Вероятность появления m и n раз противоположных событий E и F в очень большом числе $(m + n)$ испытаний. Снижение этой вероятности, когда постоянные шансы E и F, не известные заранее, выводятся из большого числа наблюдений. Пример особого случая, когда шансы этих двух событий изменяются во время испытаний (§§ 69 – 72)

Преобразование части биномиальной формулы в другую формулу, обратимую в определенный интеграл. Применение к нему метода Лапласа. Формулы, которые определяют вероятность, что в $(m + n)$ испытаниях событие E произойдет по меньшей мере m раз, а противоположное событие F – не более n раз. Вероятность, что эти числа m и n содержатся внутри [определенных] границ почти пропорциональных соответствующим шансам этих двух событий. Вероятности, что одно из этих чисел не достигнет одной или другой из этих двух границ (§§ 73 – 79)

Предыдущие формулы приводят к теореме Якоба Бернулли, изложенной в § 49. Случай, когда шанс одного из двух событий E и F очень низок. Вероятности разности между m и n содержаться внутри данных границ, когда шансы E и F либо равны друг другу, либо нет. Роль случая при очень большом числе испытаний $(m + n)$ (§§ 80 – 82)

Вероятность неизвестному шансу события E содержаться в [указанных] пределах по заданному числу раз его появления в очень большом числе испытаний. Бесконечно малая вероятность, что этот шанс в точности равен данной дроби. Мы выводим отсюда вероятность будущего события, состоящего из E и из противоположного события F. Приложение формулы, к которой мы приходим, к различным примерам. Вероятность, что событие E, происшедшее m раз в $(m + n)$ испытаниях, будет иметь место m_1 раз при другом очень большом числе испытаний $(m_1 + n_1)$. Выражение этой вероятности, соответствующей данной разности между соотношениями $m/(m + n)$ и $m_1/(m_1 + n_1)$. Сравнение шансов двух различных событий, которые появились данное число раз в заданном числе испытаний. Численное приложение предыдущих формул к примеру § 50, взятому из работ Бюффона (§§ 83 – 89)

Решение задачи, которая может иметь важное применение. Следствия о выборах депутатов очень большим числом избирателей, распределенным среди значительного количества участков, в каждом из которых происходят выборы одного депутата (§§ 90 – 93)

Гл. 4-я. Продолжение исчисления вероятностей, зависящих от очень больших чисел; случай как угодно переменных

шансов, включающий случай постоянных шансов.
Преобразование правила § 20 в формулу, выраженную определенным интегралом. Ее применение в случае очень большого числа испытаний. Определение заданных пределов вероятности того, что в $(m + n)$ испытаниях событие E произошло m раз. Отсюда мы заключаем в соответствии с первым общим предложением § 52, что это число m весьма вероятно почти пропорционально среднему из шансов E в этой серии испытаний (§§ 94 – 96)

Вероятность, что сумма значений некоторой вещи, происшедшей при заданном числе испытаний, находится внутри заданных пределов, когда число [ее] возможных значений либо ограничено, либо становится бесконечным. В частном случае, когда все эти значения имеют один и тот же шанс, постоянный в течение испытаний, эта вероятность выражается в замкнутой форме через определенные интегралы. Проверка частного результата и общей формулы в простейшем случае одного-единственного испытания (§§ 97 – 100)

Применяя эту формулу в случае большого числа наблюдений, мы доказываем теорему § 53. В соответствии с ней, если это число возрастает все больше и больше, среднее из значений вещи, которую мы рассматриваем, также приближается к постоянному значению k, с которым она совпадает, если число наблюдений может стать бесконечным. Эта особая постоянная зависит от закона вероятностей всех возможных значений [вещи]. Более или менее вероятные пределы разности δ этой постоянной и среднего из наблюденных значений при очень большом числе наблюдений зависят также от другой постоянной h этого же закона. Определение этих двух величин, k и h, в простейшем предположении о законе вероятностей. Исследование случая, когда, при реализации этого закона, число возможных значений ограничено (§§ 101 – 103)

Доказательство второго общего предложения § 52[5]. Оно завершает априорное доказательство всеобъемлющего закона больших чисел, который до сих пор считался опытным фактом (§ 104)

Правило для определения результата наблюдений, пределов разности δ, которые обладают заданной вероятностью, и, обратно, вероятности, которая соответствует заданной величине этих пределов (§§ 105 – 106)

Вероятность заданных пределов разности между средними значениями одной и той же вещи, полученными из двух различных серий испытаний. Наиболее предпочтительное правило, чтобы из двух или более серий наблюдений заключить приближенное значение этой вещи в случае, когда средние значения действительно стремятся к ее точному значению, т. е. в случае, когда специальная постоянная в каждой серии является этим истинным значением (§§ 107 – 108)

Вероятность заданных пределов разности между отношениями $m/(m + n)$ и $m_1/(m_1 + n_1)$ количеств раз m и m_1 появления одного и того же события E в $(m + n)$ и $(m_1 + n_1)$ испытаниях, когда все возможные причины E в обеих сериях наблюдений совпадают, хотя шансы этого события как-то изменяются в течение каждой серии (§ 109)

Решение задачи о наклонностях планетных орбит к эклиптике и их эксцентриситетах. Решение подобной задачи о наклонностях комет[ных орбит]. Мы заключаем с очень высокой вероятностью, что неизвестная причина образования комет не привела к неравным вероятностям ни их различных наклонностей к эклиптике, ни направлений их движения, прямого или попятного. Кроме того, оказалось, что средняя наклонность всех существующих комет вероятно очень мало отличается от той же величины для комет, известных до сего дня. Примечание относительно раскаленных тел, которые в очень большом числе наблюдаются в определенное время года[6] (§§ 110 – 111)

Таблица наиболее применяемых формул вероятностей, доказанных в этой и в предыдущих главах. Замечание о приложении исчисления вероятностей к системе *условных уравнений*, доставляемых наблюдениями (§§ 112 – 113)

Гл. 5-я. **Приложение общих правил вероятностей к решениям судов присяжных и суждениям трибуналов.**

Определение вероятностей, что обвиняемый будет осужден или оправдан определенным большинством присяжных, каждый из которых имеет заданную вероятность не ошибаться, с учетом также заданной вероятности виновности обвиняемого, которая имеет место до вынесения судебного решения. По правилам вероятностей причин и гипотез определяются также вероятности, что обвиняемый, таким образом осужденный или оправданный, виновен или нет (§§ 114 – 117)

Формулы для случая некоторого числа присяжных, имеющих один и тот же шанс не ошибаться и выносящих решения либо заданным большинством, либо большинством, для которого дана лишь его наименьшая допустимая величина. Указано, что вероятность осуждения всегда ниже вероятности виновности до судебного решения. При прочих равных условиях вероятность добротного решения зависит лишь от большинства голосов, при котором оно было вынесено, но никак не от общего числа присяжных[7], если шанс их ошибки известен заранее, а не выведен апостериорно по известному большинству (§§ 118 –120)

Приложение этих формул к случаю очень большого числа присяжных, что делает очень маловероятным, что осуждение выносится слабым большинством (§ 121)

Теорема о суде, состоящем из какого-то числа присяжных, каждый из которых имеет многочисленные и не равновероятные шансы не ошибаться. Пример вычисления среднего шанса, когда число всех возможных шансов становится бесконечным и закон их вероятности задан. Этот средний шанс один и тот же для всех присяжных, если они должны быть выбраны случайно из одного и того же общего списка. Формулы, которые в этом случае определяют вероятности осуждения; виновности осужденного и шансу ошибки присяжных находиться в заданных пределах (§§ 122 – 127)

Приложение этих формул к суду, состоящему из очень большого числа присяжных (§§ 128 – 131)

Во всех случаях применение этих формул требует принятия предположения о законе вероятности шансов ошибок присяжных. Исследование предположения Лапласа. Следствия, которые вытекают из него и делают его неприемлемым. Невозможность принятия при этом [при заранее принятом] законе ни какого-либо надлежаще обоснованного предположения, ни определения вероятности добротности отдельного судебного решения при знании числа присяжных и большинства, при котором оно было вынесено. Необходимость прибегать к результатам очень большого числа судебных решений, чтобы вывести два специальных элемента, включенных в предыдущие формулы, а именно шанс u не ошибаться, общий для всех присяжных, выбранных случайно из одного и того же общего списка, и вероятность k вины обвиняемого, выводимая из судопроизводства, которое предшествовало прениям сторон в суде (§§ 132 – 133)

Вероятность, что разность между долей осуждений, доставленным серией испытаний, и специальным значением, которого она достигает, если числа обвиняемых и осуждённых становятся бесконечными, заключена между заданными пределами. Вероятность, что разность между первой долей и той, которая доставляется другой серией испытаний, также заключена между заданными пределами (§ 134)

Из *Compte généraux de l'administration de la justice criminelle* мы выбрали данные наблюдений, которые служат для определения численных значений *u* и *k*. Эти данные являются различными отношениями, которые применяются перед использованием предыдущих формул вероятностей. Влияние последовательных изменений законодательства о присяжных во Франции на величину этих отношений. Разделение преступлений на две различающиеся друг от друга категории. Необходимое в настоящее время предположение, что значения *u* и *k*, весьма различные для этих двух категорий, почти одни и те же для всех департаментов (§§ 135 – 138)

Вычисление этих значений для Франции в целом и для департамента Сена в частности. Вероятность, что в соответствии с этими значениями решение об осуждении или оправдании вынесено единогласно (§§ 139 – 141)

Смысл, который следует приписать словам *виновный* и *невиновный*. Это более подробно пояснено в Предисловии (§ 142)

Формулы, которые указывают меру опасности для обвиняемого быть осужденным вопреки должному порядку, и для общества от оправдания обвиняемого, который должен был бы быть осужден (§ 143)

Вычисление численных значений этих мер и вероятностей невиновности и виновности осужденных в различные годы, когда законодательство не оставалось тем же самым (§§ 144 – 145)

Указание подобного вычисления, которое не может быть выполнено ввиду отсутствия необходимых данных наблюдения и которое относится к решениям о незначительных правонарушениях и военной юстиции (§ 146)

Формулы о более или менее вероятной добротности судебных решений в гражданских делах, вынесенных в первой инстанции и при апелляции (§§ 147 – 149)

Отсутствие данных опыта, необходимого для определения двух различных величин, которые входят в эти формулы, делает необходимым предположение о равенстве шансов ошибок для всех судей обеих последовательных инстанций. Вычисление этих шансов по отношению, доставленному опытом, числа судебных решений, подтвержденных королевскими судами, к числу решений судов первой инстанции [по делам,] которые представляются им на рассмотрение ежегодно. Малые изменения этого отношения в течение трех последовательных лет является весьма примечательным доказательством общего закона больших чисел. Из этих данных наблюдения выводятся вероятности добротности судебных решений первой инстанции и апелляционных судов при подтверждении этих решений, и в противном случае (§§ 150 – 151)

Примечания

1. Ввиду ошибочной нумерации пропущен § 12, так что после § 11 следует § 13. Автор

2. Словом *вещь A* Пуассон неизменно обозначал случайную величину. Пуассон ещё до 1837 г. (1829, с. 3) ввёл случайную величину формально, хотя и назвал её явно временным термином, и притом в теорию вероятностей вообще. Буква А в его термине вряд ли имела отношение к *aléa* (случайность, риск), поскольку он (1811/1833, с. 141, 146) той же буквой обозначил константу.

3. Не по правилам, а по определению.

4. В § 2 Пуассон лишь упомянул, что возможны и иррациональные значения вероятности.

5. Применяя это общее предложение, например, к *терапевтике*, мы получаем то, что к тому же соответствует простому здравому смыслу. Если лекарство было успешно применено в очень большом числе схожих случаев, так что количество неудач было очень невелико по сравнению с общим числом опытов, весьма вероятно, что оно окажется успешным и при новом испытании. Медицина не станет ни наукой, ни искусством, если не будет основываться на многочисленных наблюдениях, а также на такте и должном опыте врача, чтобы судить о сходности случаев и оценивать исключительные обстоятельства. Автор

6. Представляется, что эти тела во время их появления очень удалены от Земли и находятся на расстоянии, на котором плотность атмосферы никак не ощутима, так что трудно приписать, как это [все-таки] принято делать, их раскаленность трению о молекулы воздуха. Нельзя ли предположить, что электрические флюиды, находящиеся в нейтральном состоянии, образуют своего рода атмосферу, которая простирается намного выше воздушной массы и, хоть и является физически невесомой, подвержена притяжению Земли и потому следует в своих движениях за земным шаром? В соответствии с этим предположением, тела, о которых идет речь, и вообще *аэролиты* [каменные метеориты], входя в эту невесомую атмосферу, разлагают нейтральные флюиды своим неравным воздействием на электричества [на электрические заряды] двух родов и нагреваются и раскаливаются ввиду электризации. Автор

7. Это заметил и Остроградский (1836), который зачитал свой доклад в 1834 г. См. также Прим. 1 к гл. 5.

Библиография

Остроградский М. В. (1836, франц.), Извлечение из мемуара о вероятности судебных ошибок. *Полн. собр. трудов*, т. 3. Киев, 1961, с. 65 – 70.

Poisson S. D. (1811), *Traité de mécanique*, t. 1. Paris, 1833.

--- (1829), Sur la probabilité des résultats moyens des observations. *Conn. des temps* pour 1832, pp. 3 – 22.

Предисловие

[1] Задача по поводу азартных игр, которую светский человек предложил строгому янсенисту[1], оказалась началом исчисления вероятностей. Она имела целью установить пропорцию, в которой *ставку* следует разделить между игроками в том случае, когда они соглашаются не продолжать игру, а принять за основание победы в ней неравные количества [недостающих] очков. Паскаль первым решил эту задачу, но лишь для случая двух игроков, а затем её решил Ферма в общем случае любого числа игроков. И все-таки геометры XVII в., которые занимались исчислением вероятностей, лишь определяли шансы в различных играх той эпохи[2], и только в следующем столетии это исчисление развилось и стало одной из основных ветвей математики и по числу и пользе своих приложений, и по характеру анализа, которое оно породило.

Среди его приложений одно из наиболее важных относится к вероятности [справедливости] приговоров, или, вообще, решений, принимаемых большинством голосов. Первым, кто попытался её исследовать, был Кондорсе. Книга, которую он написал (Condorcet 1785), была предпринята при жизни и по предложению министра [финансов 1774 – 1776 гг. А. Р. Ж.] Тюрго [1727 – 1781], представлявшего себе всю пользу, которую моральные науки и гражданская администрация могут извлечь из исчисления вероятностей, чьи указания были всегда точны даже когда они, ввиду недостаточности наблюдений, не могли полностью решить поставленные задачи.

В своем предварительном [но] весьма обширном сочинении Кондорсе без помощи аналитических формул излагает полученные им результаты и тщательно развивает соображения, подходящие для того, чтобы выявить пользу подобных исследований. В своем *Трактате о вероятностях* [*Аналитической теории вероятностей*] Лаплас также занимался вычислением шансов ошибки, которой следует опасаться в приговорах, выносимых обвиняемому известным большинством голосов. Данное им решение этой задачи, одно из самых утонченных в теории вероятностей[3], основано на принципе, служащем для определения вероятностей различных причин, которые могут быть приписаны наблюденным событиям. Бейес вначале сформулировал этот принцип в несколько ином виде, Лаплас же и в своих [ранних] мемуарах, и в своем трактате, нашел ему затем удачнейшее приложение для вычисления вероятностей будущих событий по наблюдению прошлых. Однако, что касается задачи о вероятности приговоров, то справедливости ради следует сказать, что именно Кондорсе пришла в голову находчивая мысль обосновать её решение на принципе Бейеса при последовательном рассмотрении вины и невиновности обвиняемого как неизвестной причины вынесенного приговора. Иными словами, обосновать ее рассмотрением наблюденного факта, из которого выводится вероятность этой причины. Точность принципа Бейеса наглядно рассмотрена со всей строгостью, и по приложению к вопросу, которым мы занимаемся, не может больше оставлять никакого сомнения[4].

Тем не менее, Лаплас допустил при этом предположение, которое никак не назовешь неоспоримым. Он решил, что вероятность непогрешимости присяжного заседателя может принимать все равновозможные степени от достоверности, представленной единицей, до безразличия, которой в вычислениях соответствует дробь 1/2, придающей равные шансы ошибке и истине. Прославленный геометр[5] обосновал свое предположение тем, что мнение присяжного несомненно склоняется к истине, а не к ошибочности, и с этим следует в общем действительно согласиться. Но существует бесконечное множество различных законов вероятностей ошибок, которые удовлетворяют этому условию, т. е. требуют, чтобы шанс безошибочности присяжного никогда не опускался ниже 1/2, а был выше этого предела, без того, чтобы все его значения были равновозможными. Это частное предположение Лапласа поэтому не обосновано априорно. Либо по этой причине, либо ввиду её последствий, которые мне представляются неприемлемыми, отличающиеся друг от друга решения задачи о вероятности [справедливости] приговоров, содержащиеся в *Трактате о вероятностях*, с. 460, и в Первом Приложении (с. 32) к этому великому сочинению [с. 469 – 470 и 528 в издании 1886 г.], постоянно оставляют в моем сознании сильное сомнение.

[2] Я сообщил бы об этом прославленному автору, занимайся я этой задачей при его жизни. Авторитет его имени сделал бы это моим долгом, а его дружба, которая всегда возвышала меня, облегчила бы мне это. Можно без труда представить себе, что лишь после длительного раздумья я решился рассматривать упомянутый вопрос с другой точки зрения. И прежде чем идти дальше, мне будет позволительно указать основные причины, побудившие меня отбросить последнее решение, на котором остановился Лаплас и которое он дополнил численными результатами в *Опыте* [1814, в главе о вероятности судебных приговоров].

Формула Лапласа, выражающая вероятность ошибки в приговоре, зависит лишь от большинства голосов, которое его вынесло, и от общего числа судей, но не содержит ничего относительно их более или менее обширного знания порученного им дела. Из этого следует, что вероятность ошибки в решении, вынесенном судом присяжных, большинством, например, в семь голосов против пяти, окажется одной и той же, каково бы ни было множество лиц, из которых выбраны эти 12 присяжных. И это следствие показалось мне достаточным, чтобы никак не принимать ту формулу, из которой оно было выведено.

Та же самая формула предполагает, что до решения суда не существовала никакая презумпция вины обвиняемого [существовала презумпция его невиновности], так что более или менее высокая вероятность его вины должна следовать только из этого решения, вынесенного против него. Но это снова недопустимо. Перед тем, как обвиняемый предстанет перед ассизами[6], он находится в предварительном заключении[7], что [уже] установило вероятность, более высокую, чем 1/2, что он виновен. И, конечно же, каждый, не задумываясь, стал бы держать пари в справедливой игре, что обвиняемый скорее виновен, чем нет.

Между тем, правила для установления вероятности наблюденного события по вероятности причины, которые являются основанием занимающей нас теории, требуют принимать во внимание все предшествовавшие предположения, если только они не считаются отсутствующими, или если доказано, что их нет. Но подобное предположение очевидно по [имевшему место] судопроизводству, и я должен учитывать его при решении поставленной задачи. И по существу ясно, что в противном случае будет невозможно согласовать последствия вычислений с неизменными результатами наблюдений. Это предположение подобно тому, которое имеет место в приговорах по гражданским делам, когда один из судящихся подает апелляцию на первичное судебное решение в суд высшей инстанции: он появляется там вместе с презумпцией, противоречащей его делу. И без учета этого обстоятельства подсчет вероятности ошибки, которой можно опасаться в окончательном приговоре, завершится серьезным заблуждением.

Наконец, Лаплас ограничился рассмотрением вероятности ошибки приговора, вынесенного заданным большинством голосов, однако опасность для обвиняемого, преданного суду, быть ошибочно признанным виновным этим большинством, зависит не только от этой вероятности. Она также зависит от шанса, что подобный приговор будет вынесен. Итак, условно принимая, что вероятность ошибки в приговоре, вынесенном большинством семи голосов против пяти, выражается дробью почти равной 2/7, что следует из формулы Лапласа, необходимо также заметить, что в соответствии с опытом[8] число ежегодных осуждений судом присяжных во Франции указанным большинством составляет лишь 7/100 общего числа обвиняемых. И опасность для обвиняемого быть неправедно осужденным таким большинством следует измерять произведением двух дробей, 2/7 и 7/100, равным 1/50, ибо во всех случайных вещах опасение потери или надежда на выгоду выражается произведением стоимости вещи, которой опасаются или на которую надеются, на вероятность, что она произойдет.

Это рассуждение уже сводит к одной пятидесятой долю невиновных обвиняемых, ежегодно осуждаемых наименьшим возможным большинством голосов. И это еще без сомнения слишком много, если только все эти обвиняемые действительно невиновны. Но именно здесь удобно пояснить истинный смысл, который надлежит приписывать в этой теории словам *виновный* и *невиновный,* и который Лаплас и Кондорсе [Кондорсе и Лаплас] по существу им и приписывали. Никогда нельзя придти к математической проверке вины обвиняемого. Даже его признание можно считать лишь [её] вероятностью, весьма близкой к достоверности. И поэтому наиболее просвещенные и человечные присяжные заседатели выносят обвинительный приговор лишь при высокой вероятности [вины], часто, однако, низшей, чем та, которая следует из признания обвиняемого.

Между этими присяжными и судьями в гражданских исках есть существенное различие. Если судья, после глубокого изучения процесса, может ввиду сложности вопроса выявить лишь слабую вероятность в пользу одной из сторон, этого достаточно, чтобы обвинить противную сторону [чтобы решить против нее]. Однако, присяжный должен подавать свой голос за обвинение лишь если в его глазах вероятность, что обвиняемый виновен, достигла определенной границы и намного превышает вероятность его невиновности.

[3] Как ни старайся, всех шансов ошибки в уголовных делах избежать нельзя, так в какой же степени они должны быть снижены, чтобы обеспечить невиновности наибольшую возможную гарантию? На этот вопрос трудно дать общий ответ. По Кондорсе, шанс быть неправедно осужденным должен быть равносилен шансу такой опасности, которую мы полагаем достаточно малой, чтобы не стараться избегать ее в обычной жизни. Ибо, говорит он, общество вполне имеет право для [своей] безопасности подвергать своего члена риску, шанс которого для него, так сказать, безразличен. Но это рассуждение чересчур утонченно для столь серьезного вопроса.

Лаплас привел определение, намного более подходящее для освещения вопроса о шансе ошибки, которую мы вынуждены принимать в судебных решениях по уголовным делам. Он полагал, что эта вероятность должна быть такой, чтобы риск для общественной безопасности оказался выше при оправдании виновного, чем страх обвинить невинного. Как он недвусмысленно сказал, каждый присяжный призван решить в соответствии со своим методом, просвещением и мнением, скорее именно этот [подразумеваемый] вопрос, чем определить виновность обвиняемого. И ошибка при его голосовании, когда он либо осуждает, либо оправдывает, может происходить от двух различных причин: либо он скверно оценивает доводы за и против обвиняемого, либо устанавливает слишком высокую или слишком низкую границу вероятности, необходимую для осуждения. Эта граница не одна и та же для всех лиц, призванных выносить приговоры, она меняется и вместе с сутью обвинений, и даже зависит от обстоятельств, при которых происходит суд. В армии, перед лицом неприятеля, и при рассмотрении дел о шпионаже, эта граница безусловно намного ниже, чем в обычных случаях. Она снижается, а число осуждений возрастает для тех видов преступлений, которые учащаются и опаснее для общества.

Решения суда присяжных относятся к уместности осуждений или оправданий. Язык уточнится при употреблении слов *подлежащий осуждению*, в которых вся правда, взамен *виновен*, которое требует разъяснения и которое мы продолжаем применять, чтобы соответствовать обычаю. Итак, когда мы устанавливаем, что среди очень большого числа судебных решений имеется определенная доля неправедных осуждений, не следует понимать, что она состоит из невинно осужденных. Эта та доля осужденных, для которой вероятность оказалась слишком низкой не для того, чтобы установить, что они скорее виновны, чем нет, а для того, чтобы осуждение оказалось необходимым для общественной безопасности. Определение среди этих осужденных числа тех, которые действительно невиновны, не входит в задачу наших вычислений. Тем не менее, можно полагать, что это число, по крайней мере исключая политические процессы, по счастью очень незначительно. В обычных случаях об этом можно судить по очень небольшому числу осуждений, вынесенных судом присяжных, против которых восстало общественное мнение; по малому числу предоставленных [судами высших инстанций] полных оправданий; и также по очень малому числу случаев, когда ассизы, используя данное им по закону право отменять обвинительные приговоры, вынесенные судом присяжных, возвращают обвиняемого под суд других присяжных, если решат, что [проведенные] устные прения сторон разрушили обвинение и осужденный не виновен.

Результаты относительно шансов ошибки в судебных решениях по уголовным делам, которые получил Лаплас, представляются непомерными и расходящимися с общими идеями, что противоречит его словам, что *теория вероятностей – это в сущности не что иное, как здравый смысл, сведенный к исчислению* [1814/1999, с. 863 правый столбец]. Их [результаты] скверно истолковали, и было бы чересчур поспешно заключать, что математический анализ никак нельзя применять ни к вопросам такого рода, ни вообще к вещам, называемым моральными.

[4] Этот предрассудок, как я с сожалением вижу, разделяется доброжелательными людьми; чтобы его развеять, я считаю полезным напомнить здесь некоторые общие соображения, подходящие также ради сравнения их с другими вопросами, которых никто не оспаривает и в которых приложение исчисления вероятностей считается законным и необходимым. Напомнить и для того, чтобы хорошо понять назначение той задачи, которую я специально предлагаю в этом сочинении.

Вещи каждой природы подвержены всеобъемлющему закону, который можно назвать <u>законом больших чисел</u>. Он состоит в том, что, если наблюдать весьма значительное число событий одного и того же вида, зависящие и от постоянных причин, и от причин, беспорядочно изменяющихся то в одном направлении, то в другом, т. е. так, чтобы их изменения не происходили в одном каком-нибудь определенном смысле монотонно, то среди этих чисел обнаружатся почти постоянные соотношения.

Для вещей каждой природы эти соотношения имеют [свое] специальное значение, от которого они уклоняются все меньше и меньше по мере того, как ряд наблюденных событий всё удлиняется и которого они строго достигнут, если возможно будет продлить этот ряд до бесконечности. По мере того, как амплитуды вариаций беспорядочно изменяющихся причин оказываются бóльшими или меньшими, потребуется также бóльшее или меньшее число наблюденных событий, чтобы их соотношения достигли разумного постоянства. В каждом вопросе само наблюдение покажет, достаточно ли продолжен ряд испытаний. И, смотря по числу установленных фактов и величине уклонений, которые еще остаются между указанными соотношениями, вычисление представит достоверные правила для определения вероятности, что специальное значение, к которому сходятся эти соотношения, заключено в сколь угодно тесные пределы.

Если произвести новые испытания, и если обнаружится, что те же соотношения заметно уклоняются от своих окончательных значений, определяемых предыдущими наблюдениями, можно будет заключить, что в интервале между двумя рядами испытаний причины, от которых зависят наблюдения, претерпели монотонное или даже внезапное изменение. Однако, без помощи исчисления вероятностей можно серьезно ошибиться в необходимости подобного заключения. Оно же не оставляет ничего неопределенного в этом смысле и предоставляет нам также необходимые правила для установления шанса изменения причин, указанного сравнением наблюденных событий в различные эпохи.

Этот закон больших чисел наблюдается в событиях, которые мы приписываем слепому случаю ввиду незнания их причин или потому, что эти причины слишком сложны[9]. Так происходит в играх, в которых обстоятельства, определяющие появление некоторой карты или некоторого числа очков на кости, бесконечно изменяются и не могут быть подвергнуты никакому вычислению. Но если продолжить надолго ряд испытаний, различные случаи все же произойдут в соответствии с постоянными соотношениями. Более того, когда удается вычислить в соответствии с правилами какой-то игры относительные вероятности возможных случаев, подтверждается, что они равны этим постоянным соотношениям сообразно с известной теоремой Якоба Бернулли. Но в большинстве вопросов о случайности априорное определение шансов различных событий невозможно, и, напротив, они становятся известными по наблюденным результатам.

[5] Нельзя, например, заранее вычислить вероятность потери судна во время длительного плавания, и мы заменяем это вычисление сравнением числа кораблекрушений с числом плаваний. Когда последнее очень велико, отношение этих чисел почти постоянно, по крайней мере для каждого моря и каждой нации по отдельности. Его значение может быть принято за вероятность будущего кораблекрушения, и на этом естественном следствии закона больших чисел зиждется морское страхование. Если страхователь основывается лишь на незначительном числе случаев, то это простое пари, ставки в котором нельзя вычислить; если же он опирается на очень большие числа, то это рискованная операция, успех которой почти достоверен.

Тот же закон равным образом управляет явлениями, вызываемыми известными силами совместно со случайными причинами, влияние которых никак не закономерно. Последовательные повышения и понижения [уровня] моря в портах и на побережьях предоставляют пример замечательной точности. Несмотря на неравенства, вызываемые ветрами, которые устраняют законы этого явления в отдельных или малочисленных наблюдениях, приняв среднее из большого числа наблюденных морских приливов и отливов в одном и том же месте, мы обнаруживаем, что они почти соответствуют законам *приливов и отливов*, вызванных притяжением Луны и Солнца и совпадали бы с ними, будто нет никакого влияния случайных ветров. То действие, которые ветры, дующие в одном и том же направлении в течение части года, могут оказать в этот период на море, еще нисколько не определено. Различие в указанных средних, выведенное из наблюдений в начале и конце прошлого века и отделенных друг от друга на сто лет, было лишь небольшим и может быть приписано некоторому изменению, происшедшему в окрестностях.

В качестве примера закона, который я рассматриваю, укажу еще на длительность среднего срока жизни рода человеческого. Из значительного числа детей, рожденных в достаточно близких местах и эпохах, есть умирающие в младенческом возрасте, другие живут дольше, а некоторые доживают до границы продолжительности жизни. И, несмотря на превратности жизни человека, которые приводят к столь большим различиям в возрастах умирающих, если разделить сумму этих возрастов на число людей, предположенное очень большим, то частное, или то, что называется *средним сроком жизни*, будет величиной, не зависящей от этого числа. Этот срок может не быть одним и тем же для обоих полов[10], быть различным в разных странах и в различные эпохи, потому что он зависит от климата и без сомнения также от благосостояния людей. Он возрастает, если какое-нибудь заболевание исчезает, как оспа вследствие благодеяния вакцины. И во всех случаях исчисление вероятностей нам указывает, что если выявленные изменения в этом сроке довольно велики и относятся к достаточно большому числу наблюдений, то их необходимо приписать каким-то изменениям, происшедшим в общих причинах.

Соотношение между ежегодными мужскими и женскими рождениями в обширном государстве также постоянно и, видимо, не зависит от климата, но ввиду какой-то особенности, которой быть может нетрудно приписать правдоподобную причину, представляется различным для законнорожденных детей и тех, кто родился вне брака.

Строение тел, образованных разобщенными молекулами, разделенными друг от друга пространствами, свободными от весомой материи, также являет нам особого рода приложение закона больших чисел. Если из точки, выбранной внутри тела, провести прямую в определенном направлении до первой встреченной молекулы, то полученное расстояние, хоть и мало в любом смысле, будет тем не менее изменяться в очень большом отношении [в очень широких пределах] с направлением. Оно может быть в 10, 20, 100, ... раз больше в одном направлении, чем в другом. Распределение молекул вокруг каждой точки может быть очень беспорядочным и весьма различным в различных точках. Оно непрерывно изменяется ввиду внутренних колебаний молекул, потому что тело, находящееся в покое, это просто множество молекул, которые совершают непрерывные колебания с незаметными, но сравнимыми с расстояниями между ними амплитудами. И если разделить каждый кусок объема неощутимой величины на число содержащихся в нем молекул, которое исключительно велико по причине их непомерной малости, и извлечь из частного кубический корень, мы получим *средний интервал* между молекулами, не зависящий от беспорядочности их распределения. Он постоянен по всему протяжению однородного тела, имеющего повсюду одну и ту же температуру, если только отвлечься от неравного сжатия его частей, вызванного его собственным весом. И на подобных рассуждениях основано вычисление молекулярных сил и теплового излучения внутри тел, подобное тому, которое я представил в других сочинениях[11].

Все эти примеры закона больших чисел взяты из разряда физических вещей, и, если потребуется, мы можем их еще больше умножить. И не труднее указать другие примеры, относящиеся к вещам из разряда моральных. Среди них мы можем назвать доходы от косвенных налогов, постоянные если не ежегодно, то по крайней мере в течение малого числа последовательных лет. Такова, среди прочих, *судебная пошлина*, которая ежегодно прибавляет почти постоянную сумму в доход государства, но тем не менее зависит от числа и значимости дел, т. е. от противоположных и переменных интересов граждан и их большей или меньшей способности вести свои дела в суде.

Таковыми были также доходы от лотереи Франции до тех пор, пока ее, к счастью, не прикрыли, и игры Парижа, упразднение которых не менее желательно[12]. Эти игры предоставляют постоянные соотношения двух различных видов: с одной стороны, сумма ставок почти та же каждый год или в течение каждого периода небольшого числа лет. С другой стороны, прибыль банка ощутимо [в достаточной степени] пропорциональна этой сумме. Итак, эта пропорциональность является естественным следствием случая, который в соответствии с правилами игры отдает банку благоприятные исходы в постоянной и заранее вычислимой пропорции. Но постоянство суммы ставок это событие, которое относится к моральному разряду, потому что поставленные деньги зависят от числа и желания игроков. Хорошо, что эти два элемента, пропорция дохода и сумма ставок, мало изменяются, без чего откупщик игры не смог бы оценить наперед ежегодную плату, которую он будет обязан уплатить правительству, по своей прибыли в период предшествовавшего арендного договора.

[6] Ниже я изложу данные опыта, на которые я опираюсь в вопросе о вероятности [справедливости] приговоров и предоставляют дополнительные решающие примеры закона больших чисел, наблюдаемые в разряде моральных вещей. Будет видно, что при действии одного и того же законодательства доля осужденных по всей Франции очень мало изменяется от года к году, так что достаточно рассмотреть примерно 7000 случаев, т. е. чисел судебных решений, выносимых каждый год судами присяжных, чтобы эта доля ощутимо достигла постоянства, тогда как в других вопросах, как, например, о средних сроках жизни, которые я только что приводил, подобное число было бы далеко не достаточно, чтобы довести результат до постоянства. Разительно видно также влияние общих причин на рассматриваемую долю, которая каждый раз изменяется с изменением законодательства.

Нельзя, стало быть, сомневаться, что закон больших чисел соответствует моральным вещам, которые зависят от желания, интересов, просвещения и страстей человека так же, как вещам физического разряда. И по существу здесь дело совсем не в природе причин, но в изменении их отдельных влияний и числа необходимых случаев, чтобы беспорядочность наблюдаемых событий уравнивалась в средних результатах. Величина этих чисел не может быть установлена заранее, она различна в различных вопросах, и, как сказано выше, она тем значительнее, чем, в общем, больше амплитуда этой беспорядочности. Но в этом отношении нельзя считать, что влияния внезапного желания, ослепления страстями, недостатка просвещения изменяются в более широких пределах, чем человеческая жизнь, границы которой определяются младенцами, умирающими при рождении, и людьми, доживающими до ста лет. Предвидеть их труднее, чем обстоятельства, которые приводят к потери судна в длительном плавании; они более капризны, чем случай, который определяет [полученную игроком] карту или исход выбрасывания кости. Не этими идеями [влияниями?] мы объясняем действия и их причины, а вычислениями и наблюдениями, которые одни только и могут установить вероятные пределы их изменений при очень большом числе испытаний.

Из этих примеров всевозможной природы следует, что всеобъемлющий закон больших чисел уже является для нас общим и неоспоримым фактом, который вытекает из опыта и никогда не будет опровергнут. Этот закон к тому же является основанием всех приложений исчисления вероятностей. Отныне он представляет события независимо от природы вопроса совершенно одинаково, идет ли речь о физических или моральных вещах, лишь бы наблюдения обеспечили соответствующие данные, которых требует вычисление в каждой задаче. Но, имея в виду его значимость, этот закон необходимо было непосредственно доказать. И это то, что я постарался сделать, и я верю, что в конце концов достиг своей цели, как усматривается здесь ниже.

Упомянутая выше теорема Якоба Бернулли совпадает с этим законом больших чисел в том частном случае, когда шансы событий остаются постоянными в течение ряда испытаний. Это по существу и предполагает доказательство, данное автором, который, как известно, 20 лет размышлял [о ней]. Его теорема, однако, недостаточна для проблем, относящихся к повторению моральных вещей или физических явлений, шансы которых, вообще говоря, непрерывно изменяются, притом чаще всего без какой-либо закономерности. И, чтобы дополнить ее, следовало изучить вопрос с более общей точки зрения и всестороннее, но состояние математического анализа в эпоху Якоба Бернулли такой возможности не давало.

[7] Когда рассматриваешь это постоянство отношений, которое устанавливается и сохраняется между количествами появления события и очень большим числом испытаний, и это несмотря на изменения шанса указанного события в течение рассматриваемого времени, то пытаешься приписать эту столь примечательную закономерность действию некоей сокровенной и беспрестанно действующей причины. Но теория вероятностей устанавливает, что постоянство этих отношений является естественным состоянием вещей в физическом и моральном разрядах, которое само по себе поддерживается без содействия какой-либо посторонней причины. Лишь вмешательство некоей подобной причины может, напротив, воспрепятствовать ему или нарушить его.

Правительство опубликовало отчеты *Comptes généraux de l'administration de la justice criminelle* за девять минувших лет, с 1825 по 1833 гг. Именно в этом достоверно представленном с замечательной тщательностью сборнике я почерпнул все использованные документы[13]. Число ежегодных дел, рассмотренных ассизами королевства, составляло примерно 5000, а число обвиняемых – примерно 7000. С 1825 по 1830 гг. включительно уголовное законодательство не менялось, и осуждения судами присяжных выносились большинством, начиная с семи голосов против пяти, исключая случаи вмешательства суда при указанном наименьшем большинстве. В 1831 г. подобные вмешательства были отменены и установлено наименьшее большинство в восемь голосов против четырех, что должно было привести к более частым оправданиям. Доля оправдываемых в течение шести первых лет [1825 – 1830], если пренебречь тысячными долями, была равна 0.39; в течение одного-единственного года она опустилась до 0.38, а в другой раз поднялась до 0.40. Из этого следует, что в указанном периоде она изменялась от одного года к другому в одну или другую сторону лишь на сотую долю своего среднего значения[14].

Для законодательства, которое действовало до 1831 г., можно поэтому принять 0.39 за значение этой доли и считать 0.61 долей осуждений. В этот же период доля осуждений, вынесенных наименьшим большинством семи голосов против пяти, к общему числу обвиняемых, составило 0.07, и она также очень мало изменялась от года к году. Вычитая эту дробь из 0.61, получим 0.54 для доли осуждений, которые были вынесены более сильным большинством. Доля оправданий была поэтому равна 0.46, если для осуждения требовалось большинство не меньшее восьми голосов против четырех. И так по существу и случилось в течение 1831 г., ибо тогда разность между этой долей, выведенной по прошлым годам, и её наблюденным значением едва составила полу-тысячную.

В 1832 г., когда сохранялось то же самое наименьшее большинство, что и в 1831 г., закон предписал [учитывать] вопрос о *смягчающих обстоятельствах* и сокращать наказание в положительных случаях[15]. Влияние этой меры должно было облегчить осуждения судами присяжных, но в какой степени? Один лишь опыт может ответить нам, это нельзя вычислить заранее в отличие от повышения числа оправданий, которое произошло при изменении наименьшего возможного большинства голосов. Опыт показал, что в 1832 г. доля оправданий снизилась до 0.41 и оставалась почти до одной тысячной той же самой в 1833 г., в котором законодательство не изменилось. Доля осуждений до, в течение и после 1831 г. составила соответственно 61/100, 54/100 и 59/100, так что, после убывания на 0.7 [на 0.07] ввиду требования усилить большинство голосов на единицу, она увеличилась лишь на 0.5 [на 0.05] под влиянием вопроса о *смягчающих обстоятельствах* на сознание присяжных.

В течение 1832 и 1833 гг. число политических процессов, представленных ассизам, было значительным. Если добавить эти процессы к числу уголовных дел, для которых доля оправданий составила 0.41, то окажется, что она повысится до 0.43. Это уже указывает на влияние рода рассматриваемых дел на число оправданий, выносимых присяжными, что отчетливо видно в *Comptes généraux*. Уголовные дела там разделены на две главные категории: деяния, имевшие целью кражу или другие посягательства на собственность, и относящиеся к посягательствам на личности. Последние составляли примерно треть первых или четверть всех преступлений. С 1825 по 1830 гг. доля оправданий составляла лишь 0.34 в первой категории, а во второй она повысилась до 0.52, т. е. превысила долю осуждений на 0.04.

Годичные значения каждой из этих двух долей изменялись всего-навсего на 0.02 в ту или иную сторону от указанных дробей 0.34 и 0.52. Следует также заметить, что число осуждений, вынесенных наименьшим большинством в семь голосов против пяти, составило лишь 0.05 числа обвиняемых в преступлениях против собственности и дошло до 0.11 в случае обвиняемых в преступлениях против личности. Таким образом, не только осуждения пропорционально более многочисленны в первом случае, чем во втором, но они в общем выносились более сильным большинством.

Эти различия могут частично зависеть от меньшей суровости присяжных, когда речь идет о преступлениях против личности, а не против собственности. Последние, как более частые, без сомнения считаются более опасными для общества, однако различия в методах суждения в этих двух категориях недостаточны, чтобы привести к такому существенному неравенству между долями оправдания, указанному опытом. И вычисления обнаруживают, что это неравенство объясняется также более сильной презумпцией виновности, которая проистекает из сведений, собранных до вынесения суждения, скорее по отношению к краже, чем касающихся иных обвиняемых.

[8] Сборник *Comptes généraux* свидетельствует также о других соотношениях, которые большие числа делают почти неизменными, но которые я никак не использую. Так, например, с 1826 г., когда начали указывать пол обвиняемых, и до 1833 г. ежегодное отношение числа женщин, отданных под суд, к общему числу обвиняемых составляло почти 0.18. Один-единственный раз оно повысилось почти до 0.20 и один-единственный раз опустилось до 0.16. И оно было постоянно выше в делах о краже, чем в преступлениях против личности, а доля оправданных также была более значительна у женщин, чем у мужчин и доходила почти до 0.43, тогда как для обоих полов вместе она составила лишь 0.39.

Но постоянство этих различных пропорций, которое наблюдается каждый год во Франции в целом, не имеет места, когда рассматриваешь только отдельные ассизы. В одном и том же департаменте и при том же законодательстве доля оправданий значительно изменяется из года в год. Это показывает, что в юрисдикции одного такого суда годичное число уголовных дел совсем недостаточно для уравнивания беспорядочности голосования присяжных, или чтобы доля оправданий оказывалась постоянной. Эта доля к тому же изменяется от одного департамента к другому, и число дел в юрисдикции каждого ассиза также недостаточно велико, чтобы можно было решить с достаточной вероятностью, что они являются областями Франции, в которых суды присяжных либо более, либо менее склонны к суровости.

Лишь в департаменте Сены число уголовных дел достаточно, чтобы отмеченная доля оправданий изменялась не очень сильно, так что её можно сравнивать с той, которая имеет место для Франции в целом. Число лиц, отдаваемых каждый год в суды ассиза, составляет в Париже примерно 800, т. е. почти девятую часть от всего королевства. С 1825 по 1830 гг. доля оправданий изменялась в пределах 0.33 и 0.40 и ее среднее значение было равно лишь 0.35, тогда как для Франции в целом она составила 0.39 или на 0.04 более. Что касается доли осуждений, вынесенных наименьшим большинством семи голосов против пяти, она также была немного меньше для Парижа и составила всего 0.065 вместо 0.07 для всей Франции, если не различать категорий преступлений.

Таковы данные, которые нам доставил опыт до нынешнего времени относительно решений ассизов. Теперь, четкая цель теории – это вычисление для суда присяжных, состоящего из определенного числа лиц, выносящего приговоры определенным большинством и имеющего очень большое количество дел, доли оправданий и осуждений, которые весьма вероятно будут иметь место, и шанс ошибки судебного решения, принятого случайно отобранными лицами из тех, кто был или будет выбран присяжным. По моему, определение шанса ошибки при осуждении или оправдании в приговоре, вынесенном в известном отдельном деле, невозможно, во всяком случае, если основываться на исчислении никак не надежных предпосылок, которые приведут к весьма различающимся результатам и почти к тем, которых желают получить в соответствии с ними. Но для защиты общества и для гарантии, которое оно должно предоставить обвиняемому, важнее всего знать не этот относительный шанс частного судебного решения, который еще надлежит определить, а тот, который относится к множеству дел, переданным ассизам за один год или за много лет и который выводится из наблюдения и вычисления.

Вероятность ошибки какого-то обвинительного решения, умноженная на шанс, что оно имеет место, является истинной мерой опасности, которой общество подвергает невиновного. Произведение шанса ошибки в оправдании и вероятности, что оно вынесено, измеряет ту же опасность, которой подвергается само общество и которая равным образом должна быть известна, потому что только мера этой опасности может оправдать возможность неправедного осуждения. В этом важном вопросе человечности и общественного порядка ничто не может заменить аналитических формул, выражающих эти различные вероятности.

Без их помощи, если дело идет об изменении числа присяжных или о сравнении двух государств, в которых оно было различным, как же узнать, где представляется бóльшая гарантия обвиняемым и обществу, если в одном суд присяжных состоит из 12 лиц и выносит приговоры большинством [не менее чем] в восемь голосов против четырех, а в другом, например, из девяти лиц, избранных из того же списка, как в первом случае, и выносящих приговоры тем или иным большинством.

Как было решено, была ли система, которая существовала во Франции до 1831 г., с судебными решениями, принимаемыми также большинством в семь голосов против пяти и вмешательством судей в случае наименьшего большинства, более полезна или менее благоприятна, чем та, которая имеет место сегодня, – с тем же большинством и учитывая влияние вопроса о *смягчающих обстоятельствах*? Исследовать это сейчас нельзя ввиду отсутствия данных наблюдения, относящихся к нынешнему времени.

[9] Формулы, которые будут определять суть вопросов, и которые включены в это сочинение, выведены без всяких предположений из общих и известных законов исчисления вероятностей. Они содержат две особые величины, зависящие от морального состояния государства, применяемого способа уголовного судопроизводства и опытности служащих судебного ведомства, которым поручено управлять им. Одна из них выражает вероятность, что присяжный, случайно выбранный из списка в юрисдикции какого-нибудь ассиза, не ошибется при голосовании, другая – вероятность, существующая перед началом прений сторон, что обвиняемый виновен. Таковы два существенных элемента в вопросе об уголовных судебных решениях. Их численные значения должны быть выведены из данных опыта, подобно постоянным, содержащимся в астрономических формулах и выведенным из наблюдений. И всё решение задачи, которая предложена в этих исследованиях, требует взаимодействия теории и опыта.

Я использовал две величины, данные наблюдениями, и число определяемых элементов тоже равно двум. Это 1) количество осуждений большинством не менее семи голосов против пяти, и 2) то же количество при наименьшем большинстве, притом оба количества делятся на общее число обвиняемых. Полученные отношения весьма различны для преступлений против личности и посягательств на собственность, и я рассматриваю эти две категории по отдельности. Они также не одинаковы в различных департаментах, но необходимость их вывода по очень большим числам заставила меня объединить по каждой из указанных категорий решения всех ассизов королевства. Значения, которые я после этого вывел для этих двух определяемых элементов, являются лишь приближенными и предполагают, что они не очень изменяются от одного департамента к другому.

Но новый закон при восстановлении наименьшего большинства при осуждении в семь голосов против пяти предписал судам присяжных указывать приговоры, вынесенные таким образом. И после этого [нововведения] нам известно по каждому департаменту довольно значительные числа осуждений, вынесенных и таким, и каким-либо иным большинством, что необходимо для определения наших двух элементов. И таким образом мы будем знать, заметно ли изменяется шанс ошибки присяжных от одного места к другому. Для департамента Сена в отдельности вычисления уже установили, что этот шанс немного ниже, чем для остальной Франции.

Вот, фактически, основные численные результаты, которые включены в наше сочинение и сводку которых мне представляется полезным привести здесь. До 1831 г. для Франции в целом вероятность присяжному не ошибиться при своем голосовании была немного выше 2/3 в случае преступлений против личности и почти равна 13/17 для преступлений против собственности. Без различия этих категорий шанс оказался немного ниже 3/4, – все это для всей Франции, – и немного выше этой дроби для департамента Сена в частности. В то же время, другой элемент уголовных судебных решений, т. е. вероятность предварительной виновности обвиняемого, не намного превышала 1/2 и находилась между 0.53 и 0.54 для всей Франции по делам о преступлениях против личности. Она несколько превышала 2/3 для преступлений против собственности, а без различия между этими категориями она почти равнялась 0.64 и повышалась примерно до 0.68 в юрисдикции парижского ассиза. Вычитая эти различные дроби из единицы, мы получаем вероятности соответственно ошибки присяжного и осуждения.

[10] Можно заметить, что предварительная вероятность виновности обвиняемого всегда выше доли осуждений. Так, например, в случае, когда эта вероятность ниже всего и превышает 1/2 всего на 3 или 4 сотых, эта доля, как сказано выше, находится примерно на 2 сотых ниже 1/2. Таков общий результат, и формулы вероятности устанавливают, что он имеет место всегда при любой опытности служащих судебного ведомства, любом шансе ошибки присяжного и большинстве, требуемом для осуждения. Следует также заметить, что эта предварительная вероятность виновности обвиняемых относится лишь к осуждению судом присяжных в соответствии с их способом судить, т. е. в соответствии с неизвестной вероятностью, которую они требуют для осуждения и которая несомненно ниже вероятности, того, что обвиняемый действительно виновен в соответствии с предварительными данными. По существу каждый, не задумываясь, поставит, например, более одного против одного за то, что обвиняемый, судимый в ассизах за преступление против личности, виновен, хотя предварительная вероятность его вины, определенная для этой категории преступлений, очень мало превышает дробь 1/2.

В 1831 г. большинство, достаточное для осуждения, изменилось, а два элемента, которые мы рассматриваем, должны были остаться без изменения. В следующие годы проблема *смягчающих обстоятельств* без сомнения повлияла на их значение, но нам за 1832 и 1833 гг. известно только отношение общего числа осуждений к числу обвиняемых, чего недостаточно для определения этих элементов, и мы не знаем, стал ли в это время шанс присяжного ошибиться выше или ниже, чем раньше. Мы сможем это узнать лишь, если введем для второго элемента предположение, которое, возможно, намного уклонится от истины. И мы также не знаем, не изменился ли вновь шанс ошибки под влиянием действующего законодательства по причине предписанного присяжным тайного голосования[16]. Когда станет возможным по достаточному числу будущих наблюдений определить и его, и предварительный шанс виновности подсудимого, мы также узнаем, повторяя вычисления для всё более дальних эпох, не изменяются ли во Франции эти два элемента постоянно в одну или в другую сторону. И это окажется важным документом о моральном состоянии нашего государства.

Несмотря на то, что судьи безусловно более сведущи в уголовных делах, их шанс не ошибиться при голосовании, как представляется, лишь немногим отличается от того же шанса для присяжных. По существу, с 1826 по 1830 гг. осуждения при большинстве, составлявшим лишь семь голосов против пяти, имели место во всей Франции 1911 раз. Ассизы, насчитывавшие в то время пять судей, были призваны вмешиваться в таких случаях и 314 раз приняли сторону меньшинства. И вычисление покажет, что они должны были присоединиться к нему примерно 282 раза, если предположить, что вероятность не ошибиться одна и та же для судей и для присяжных. И хотя эти два числа, 314 и 282, не так уж велики, чтобы можно было с очень высокой вероятностью решить, в чем это предположение отклоняется от истины, малая разность между ними дает основание для мысли о том, что разность между шансами ошибки судьи и присяжного также весьма мала. Таким образом, для присяжных этот шанс не появляется, как можно было бы предполагать, от недостатка навыка.

При прочих равных условиях ясно, что доля осуждений убывает по мере того, как от суда требуется более сильное большинство. Если, как в Англии, и для осуждения, и для оправдания должно быть единогласие 12 присяжных, и если принять в качестве значений [наших] двух элементов уголовного судопроизводства те, которые относятся ко всей Франции без различия категорий преступлений, вероятность осуждения будет мало отличаться от одной пятидесятой, а вероятность оправданий – почти вполовину ниже. Решения оказываются весьма затруднительными, по крайней мере если не происходят своего рода соглашения между присяжными и какая-то часть из них не жертвует своим мнением. Очевидно также, что без этого единогласные оправдания оказались бы более затруднительными и более редкими, чем осуждения в отношении 2:1. Лишь выбрав случайно одно дело из 22, можно ставить один против одного за то, что осуждение или оправдание окажутся единогласными[17].

[11] После судебного решения вероятность, что обвиняемый виновен, становится намного выше или ниже чем до него в соответствии с тем, осужден ли он или оправдан. Формулы в нашем сочинении указывают значение этого изменения по заданному большинству, при котором было принято судебное решение, если только оба элемента, которые включены в них, определены из наблюдений. Если для осуждения необходимо большинство по крайней мере в восемь голосов против четырех, вероятность вины осужденного немного превысит дробь 0.98 для преступлений против личности и дробь 0.998 для преступлений против собственности. Это сводит шансы ошибки осуждения к немногому менее двух сотых и двух тысячных соответственно. Если учитывать вероятность не быть оправданным, то шансы ошибочного осуждения оказываются примерно равными одной сто пятидесятой и лишь четырем десятитысячным соответственно[18]. В то же время оказывается, что вероятности невиновности оправданного обвиняемого примерно равны в этих случаях 0.72 и 0.82.

Приняв во внимание вероятность не быть осужденным, можно также установить, что шансы виновного быть оправданным примерно равны 0.18 и 0.07, так что среди очень большого числа оправданных более 1/6 и примерно 1/14 должны были бы быть осуждены.

В течение семи лет с 1825 по 1831 гг. число осужденных большинством не менее восьми голосов против четырех во всей Франции оказалось равным примерно 6000 для преступлений против личности и примерно 22 000 для преступлений против собственности. В соответствии с шансами неверного осуждения, указанными выше, можно полагать, что примерно 40 и 9 этих обвиняемых не были виновны[19]. В то же время число оправданных, но виновных, должно было составлять более чем в 50 раз больше[20], чем не осужденных обвиняемых, т. е. равняться примерно 360 ежегодно.

Но нельзя упускать из вида то значение, которое мы придаем слову *виновный* и которое объяснено выше. Из него следует, что число 18 [?] является верхним пределом действительно невиновных осужденных, тогда как 360, напротив, это нижний предел числа оправданных, хотя никак не невиновных. Этот результат вычислений далек от того, чтобы понизить уважение, которое нам следует иметь к рассматриваемому предмету или уменьшить доверие к решениям судов присяжных. Напротив, он подходит для того, чтобы воспрепятствовать всякому преувеличению ошибки, которой можно опасаться в осуждениях. По существу, их суть не позволяет проверить их по опыту. Но эти результаты имеют общее со многими другими приложениями математики, которые не в большей степени допускают проверку и достоверность которых, как и здесь, опирается лишь на строгость выводов и точность данных наблюдения.

В годы, предшествовавшие 1831 г., для всей Франции вероятность ошибки в осуждении, вынесенном наименьшим большинством в семь голосов против пяти, была примерно равна 0.16 или 0.04 для преступлений против личности и против собственности соответственно. Без различия этих категорий ее значение было 0.06. По формуле Лапласа этот шанс ошибки был одним и тем же в обоих случаях и почти впятеро выше, чем 0.06. Кроме этого следует заметить, что вмешательство суда было в то время необходимо при указанном наименьшем большинстве, так что, если судьи подтверждали решение присяжных, шанс ошибки, равный 0.06, убывал немного меньше, чем на одну сотую. Таким образом, из 1597 осуждений, которые были вынесены при наименьшем большинстве в течение пяти лет с 1826 по 1830 гг., как можно полагать, примерно 15 или 16 были ошибочны, т. е. обвиняемых, хотя и не обязательно невиновных, не следовало осудить.

[12] Отличительная суть этой новой теории вероятностей судебных решений по уголовным делам состоит, стало быть, в определении, по данным очень большого числа дел того же самого вида, шансов ошибки судей и предварительной виновности обвиняемого. Она должна быть применима ко всем видам судебных решений, –в полицейских судах, в военной юстиции, в судопроизводстве в гражданских делах, – лишь бы в каждом из них было достаточно данных для определения этих двух элементов. Он должна также быть применима к судебным решениям, которые выносились в очень большом числе особыми трибуналами в злополучную революционную эпоху. Но в этом отношении, чтобы не оставлять никаких сомнений в общности и точности теории, необходимо некоторое пояснение. Трудность, которую этот исключительный случай представляет, никак не ускользнет от тех, кто хочет со вниманием отнестись к результатам моего труда.

Подсудимый может быть осужден либо потому, что виновен, а судьи не ошиблись, либо потому, что невиновен, а судьи ошиблись. Доля осуждений не изменяется при изменении предварительной вероятности виновности обвиняемого и безошибочности голосов каждого судьи на их дополнения до единицы. Она остается прежней когда, например, эти вероятности равны 2/3 и 3/4, и когда они равны лишь 1/3 и 1/4. Прежним она оказывается и тогда, когда обе эти вероятности очень мало отличаются от достоверности или единицы, и когда они почти нули. И в этих крайних случаях число осуждений очень мало отличается от числа обвиняемых. По этой причине уравнения, которые следует решить, чтобы определить эти вероятности, всегда допускают два действительных корня, дополняющих друг друга до единицы. Тем не менее, каждое из этих решений имеет отличающее его свойство. Принимая одно из них, получим вероятность виновности осужденного более высокую, чем его невиновности; напротив, принимая другое, получим обратное. В обычных случаях следует выбирать именно первое решение, потому что неразумно предполагать, что трибуналы как правило несправедливы или что они чаще всего судят противно здравому смыслу. Но это уже не так, когда судебные решения выносятся под влиянием страстей. И тогда уже следует принимать не разумный корень уравнения, а другое решение, которое приписывает столь высокую вероятность несправедливым осуждениям.

Большая доля осуждений, вынесенных революционными трибуналами, была недостаточно обоснована доказательством юридической вины обвиняемых. Исходя из законов, которые трибуналы должны были соблюдать, мы никак не можем выделить осужденных, бывших виновными и не виновными. Следует всегда обращать внимание на то, что в этой теории несправедливость судьи и страсти прокурора считаются шансами ошибок, равно как и чрезмерное сострадание и избыток снисходительности, и что вычисления основаны на результатах голосования, каковы бы ни были причины, которые его определяли.

В полицейских судах средняя доля осуждений для Франции в целом за девять последовательных лет находилось в пределах от 0.86 до 0.85. Но этого указания недостаточно для определения ни предварительной вероятности виновности обвиняемого, ни вероятности безошибочности судьи при голосовании. Предполагая, что судебные решения выносились тремя судьями, что, видимо, обычно имеет место, нам следует также знать, какие доли осуждений выносятся единогласно и простым большинством двух голосов против одного. Эти доли, которые не указаны наблюдениями, можно дополнительно определить лишь при некоторых произвольных допущениях.

В случае военных трибуналов также нет двух данных, необходимых для определения значений двух специальных элементов, содержащихся в формулах для вычисления вероятностей. Военные суды состоят из семи судей, и осуждения не могут выноситься меньшим большинством, чем пяти против двух. Общая доля осуждений по оценке равна 2/3. Но мы не знаем этой доли ни для происшедших единогласно, ни принятых простым большинством. Не имея этих данных, нельзя точно сравнить военную юстицию с ассизами относительно шанса ошибки осуждений и оправданий, а иметь такую возможность было бы очень интересно.

[13] Когда речь идет о судебных решениях в гражданских делах, формулы вероятности включают лишь одну специальную величину, а не две, и именно ту, которая выражает шанс судьи не ошибиться при голосовании. Как мне сообщили, решения трибуналов первой инстанции выносятся, как правило, тремя судьями. Но отношение числа дел, решение по которым принято единогласно, к числу тех, которые были решены простым большинством двух голосов против одного, неизвестно, и это делает невозможным непосредственное определение шанса ошибки при голосовании. По судебным решениям, на которые поданы апелляции в королевские суды, этот шанс можно вычислить по сравнению подтвержденных и неподтвержденных приговоров, если полагать, что шанс ошибки судей обоих трибуналов один и тот же. Хотя это предположение быть может намного уклоняется от истины, я принимаю его, чтобы иметь возможность дать пример вычисления ошибки, которой можно опасаться в судебных решениях по гражданским делам. Истина или справедливость следует из решений, по необходимости единогласных, судей, у которых нет никакого шанса ошибки. Ни в каком деле эта *абсолютная справедливость* не известна. Тем не менее, по ошибочным голосованиям и решениям можно понять то, что ей противоположно, и вопрос состоит в том, чтобы определить их вероятности и, следовательно, пропорции, в которых они почти точно происходят с высокой вероятностью при достаточно большом числе дел.

В сборнике *Compte général* [разночтение!], который публикуется правительством, можно найти количество решений судов первой инстанции, подтвержденных и отмененных королевскими судами в течение последних трех месяцев 1831 г. и в 1832 – 1833 гг. Отношение первого из этих чисел к их сумме для Франции в целом почти равно 0.68. Оно не менялось от одного года к другому даже на 1/70 своего значения, так что, несмотря на разнообразие дел, которое должно было иметь место, и несомненно на неравенство в просвещении служащих судебного ведомства, достаточно примерно 8000 выносимых ежегодно приговоров, чтобы это отношение приняло почти постоянное значение. Это является еще одним весьма примечательным примером всеобъемлющего закона больших чисел. В юрисдикции королевского суда Парижа это отношение заметно выше и дошло примерно до 0.76.

Прилагая его значение для Франции в целом, и принимая 7 для числа советников каждого королевского суда, который выносит приговоры в гражданских делах, мы найдем число, немного большее 0.68, для вероятности такого советника или судьи первой инстанции, выбранного случайно по всему королевству, не ошибаться при суждении о деле, также случайно взятом из тех, которые ежегодно подаются в суды этих двух инстанций. Кроме того, возможно, что эта вероятность различна для дел, решенных в первой инстанции, на которые ни одна из сторон не подавала апелляции.

По примеру этой дроби 0.68 можно найти, пренебрегая тысячными, 0.76 для вероятности добротности судебного решения первой инстанции, 0.95 для вероятности, относящейся к апелляционному суду, когда он подтверждает суждение первой инстанции и 0.64, когда он отменяет это суждение, и, наконец, 0.75 для вероятности в случае, когда решение королевского суда подтверждается вторым королевским судом, который исходит из тех же данных, что и первый. Значения вероятности, что и трибунал первой инстанции, и первый апелляционный суд судят добротно; что трибунал – плохо, а суд – хорошо; суд – плохо, а трибунал – хорошо; и оба – плохо, примерно равны соответственно 0.649, 0.203, 0.113 и 0.035 и в сумме составляют единицу.

Вопросы, относящиеся к вероятности судебных решений, принципы и полученные ответы на которые указаны здесь, находятся в пятой, последней главе этого труда. В четырех первых главах содержатся правила и общие формулы исчисления вероятностей, что освобождает читателя от их поиска в других источниках и позволяет рассматривать некоторые другие вопросы, чуждые специальной цели этих исследований, но подходящие для их освещения исчислением вероятностей. Читатель также найдет [в гл. 3-й] решение задачи, которая указывает, как большинство избираемой ассамблеи может измениться после новых выборов совершенно или в гораздо большей степени, чем казалось бы по соотношению избирателей, распределенных по избирательным участкам и голосующих не единогласно в каждом из них.

Примечания

1. Светский человек – А.Г. Де Мере (1610 – 1685), строгий янсенист – Б. Паскаль. Янсений (1585 – 1638) – голландский богослов, см. БСЭ, 2-е изд., т. 49, 1957.

2. В XVII веке родилась политическая арифметика и, в частности, статистика населения, а также страхование жизни. Кроме того, Гюйгенс исследовал проблемы смертности, но его рукописные сочинения стали известны лишь в конце XIX века.

3. Лаплас рассматривал эту задачу в гл. 11-й своего труда и в Первом Дополнении к нему. Вопреки Лапласу, Пуассон как правило применял термин *исчисление вероятностей*, особенно в основном тексте своего сочинения. Так же поступали впоследствии Бертран, Пуанкаре и Марков. До 1900 г. последний, впрочем, пользовался в своих литографированных курсах термином *теория вероятностей*.

4. Споры о бейесовском подходе не утихли до сегодняшнего дня.

5. С 1774 г. сам Лаплас (например, 1812/1886, с. 365) отделял себя от геометров, фактически считая себя прикладным математиком.

6. Ассизы – один из видов судов. Было бы важно указать, как в них принимались решения, – судьями, присяжными, или, как мы прочли про более позднее время, и теми, и другими совместно. Но именно этого указания в чётком виде у автора нет; в различных местах книги, как представляется, имеются в виду различные формы голосования в ассизах. И вообще в гл. 5-й упоминаются суды многих инстанций, и разобраться в них без особого исследования вряд ли можно.

7. В оригинале: предварительное заключение и arrêt d'accusation (постановление о привлечении к суду в качестве обвиняемого?).

8. Пуассон неизменно писал *опыт* и не упоминал статистику.

9. Указанное пояснение случая впоследствии повторил Пуанкаре, для которого, впрочем, основным было истолкование случая схемой малое действие – серьёзное следствие.

10. Пол начали различать в таблицах смертности еще до 1832 г. (Quetelet & Smits 1832, p. 33), а Corbaux (1833) повторил это нововведение, притом непонятным образом выделил несколько социальных слоёв населения.

11. Вряд ли эти сочинения Пуассона были изучены со стохастической точки зрения.

12. В конце 1837 г. они были запрещены. В то время в Париже было 7 игорных домов (*La Grande Enc.*, t. 21, p. 152), однако указания на *игры Парижа* (jeux de Paris) мы не нашли.

13. В гл. 5-й Пуассон исходил и из других источников.

14. Автор и в дальнейшем приводил подобные доводы.

15. Об этих обстоятельствах см. гл. 5. О. Ш.

В отчете администрации уголовной юстиции за 1834 г., только что опубликованном правительством, мы находим, что в течение этого года, когда законодательство оставалось тем же, что в двух предшествовавших годах, доля осуждений повысилась до 0.60, лишь на 1/100 по сравнению с указанными годами. Правительство Бельгии по примеру нашего также публикует *Compte général de l'administration de la justice criminelle* в своем королевстве. Суды присяжных были [там] вновь учреждены в середине 1831 г., и наименьшее большинство, необходимое для осуждения, установлено в семь голосов против пяти. Доля оправданных оказалась равной 0.41, 0.40 и 0.39 в 1832, 1833 и 1834 гг. соответственно.

И примечательно, что его среднее значение, 0.40, отличалось лишь на 1/100 от того, которое имело место во Франции при том же большинстве. Перед восстановлением судов присяжных уголовные трибуналы в Бельгии состояли из пяти судей и осуждения могли выноситься простым большинством трех против двух. Доля оправданных также очень мало изменялась от года к году, но доходила примерно лишь до 0.17, т. е. менее чем до половины той величины, которая имеет место в решениях судов присяжных. Это различие больше чем вдвое между пропорциями оправдания не вызвано лишь различием в количествах судей и присяжных (5 и 12) или лишь различием в наименьшем большинстве (три против двух и семь против пяти). Оно, указанное различие, предполагает также, как это усматривается из упомянутого отчета, что для осуждения судьи требуют вероятность, заметно низшую, чем присяжные, каковы бы ни были шансы ошибки у тех и у других. Автор

16. Присяжные не могут больше отказываться от своих решений после того, как примут участие в тайном голосовании. Но есть особый случай, который может иногда представиться и на который имеет смысл указать. Два человека, Пьер и Павел, обвиняются в краже. На вопрос, виновен ли Пьер, 4 присяжных ответили *да*, трое других – *да*, пятеро других – *нет*, и обвиняемый объявлен виновным большинством семи голосов против пяти. На тот же вопрос о Павле четверо первых ответили *да*, трое других, ответивших *да* по поводу Пьера, говорят теперь *нет*, пятеро последних отвечают *да*. Павел поэтому объявлен виновным при большинстве девяти голосов против трех. А теперь задается вопрос, осуществлена ли была кража *многими*, что при положительном ответе влечет за собой более суровое наказание. В соответствии со своим первым голосованием первые четыре присяжных отвечают *да*, а восемь остальных, которые решили, что либо Пьер, либо Павел невиновен, отвечают *нет*. Решение суда присяжных, без того, чтобы противоречить их голосованию, состоит, стало быть, в том, что оба обвиняемых виновны в краже и в то же время что кража не была совершена *многими.* Автор

17. В соответствии с документами, опубликованными в Англии и, видимо, заслуживающими доверия (Porter), число обвиняемых, ежегодно предстающих перед судом присяжных, в последнее время постоянно возрастает, как и доля осуждений. Вот результаты, извлеченные из этих документов, которые можно сравнить с имеющимися в нашем государстве. Числа относятся только к Англии и Уэльсу и соответствуют трем периодам по семь лет каждый, заканчивающиеся в 1818, 1825 и 1832 гг. В течение периода в 7 лет, который закончился в 1817 г. [?], число обвиняемых не поднималось до 35 000, а доля осуждений была немного ниже 0.60. За один только 1832-й год, последний из всех этих периодов, число обвиняемых дошло до 20 829, из которых 14 947, или почти 3/4, были осуждены.

Я не знаю, возросло или убыло это число в последующие годы. Доли наиболее слабого наказания в Англии и во Франции мало отличаются друг от друга. В Таблице видно, что в Англии число [коротких] тюремных заключений составило почти 2/3 всех осуждений. В течение 1832 и 1833 гг. во Франции первое из этих чисел превысило половину второго. Кроме прочего, таблица показывает, что в последний семилетний период число казненных составило в среднем до 60 человек в год. Во Франции это число теперь вдвое меньше и не превосходит 30 в год. Автор

Таблица автора

числа в строках соответствуют периодам 1 – 3

1. Число обвиняемых 64 538; 93 718; 127 910
2. Число осуждений 41 054; 63 418; 90 240
3. Отношение (2):(1) 0.636; 0.677; 0.705
4. Приговорено к смертной казни 5802; 7770; 9729
5. Казнено 636; 579; 414
6. Приговорено к тюремному заключению
   сроком 2 года или менее 27168; 42713; 58757

18. Слово *тысячным* исправлено в тексте Пуассона (здесь и в обоих случаях ниже на с. 24) от руки на *десятитысячным*.

19. Аналогичное исправление: число 88 зачеркнуто, вместо него указано число 9. Числа ни здесь, ни чуть ниже не пояснены.

20. *Более чем в 40 раз* переправлено на *более чем в 50 раз*.

Библиография

Buffon G. L. L. (1777), Essai d'arithmétique morale. В книге автора *Oeuvr. Philos*. Paris, 1954, pp. 456 – 488.

Condorcet M. J. A. N. (1785), *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. New York, 1972.

Courbaux Fr. (1833), *On the Natural and Mathematical Laws concerning Population, Vitality and Mortality*. London.

Laplace P. S. (1812), *Théorie analytique des probabilités. Oeuvr. Compl.*, t. 7, No. 2. Paris, 1886.

Porter G. R. *Tables of the Revenue, Population, etc of the United Kingdom*, pt. 2. Пуассон не указал ни места, ни года издания. Сочинение в 13 частях, 9-ти томах, описывающее период 1820 – 1843 гг.

Quetelet A., Smits E. (1832), *Recherches sur la reproduction et la mortalité de l'homme.* Bruxelles.

Sheynin O. B. (1986), Quetelet as a statistician. *Arch. Hist. Ex. Sci.*, vol. 36, pp. 281 – 325.

Глава 1. Общие правила о вероятностях

Опечатки/ошибки, не замеченные автором

1. В § 8, с. 41 оригинала, знаки в обеих биномиальных формулах неверны.

2. В самом конце § 9, с. 43 оригинала, указано, что некая величина немного менее (словами) $1/60 \cdot 10^6$ вместо $1/60 \cdot 10^7$. Аналогичная ошибка допущена чуть ниже.

3. В § 17, с. 58 оригинала, строка 7 снизу. Вместо урны V должна быть указана урна A.

4. В § 21, с. 66 оригинала. Правая часть второго равенства в самой нижней строке должна быть *b/m*, а не *a/m.*

Исправления внесены в перевод.

1. *Вероятность* события есть наше основание полагать, что оно произойдёт или произошло. Пусть оно имело место в одном случае, в другом оно [только] возможно, хотя его вероятность для нас не изменилась, если всё в обоих столь различных случаях было одним и тем же.

Шарик извлекли из урны, содержащей некоторое количество белых и чёрных шариков в известном мне соотношении, либо же он извлекается, но его цвет ещё неизвестен. Моё основание полагать, что шарик – белый, очевидно не изменилось. Вероятность зависит от нашего знания о данном событии; она может быть различной для разных людей. Так, если некто знает лишь, что в урне находятся белые и чёрные шарики, а другому лицу известно ещё, что белых шариков больше, чем чёрных, это лицо может более обоснованно полагать, что появится белый. Иначе говоря, извлечение белого шарика оказывается для него более вероятным, чем для первого лица.

По этой причине, имея различные знания о каком-либо вероятном событии, два человека, A и B, иногда приходят к противоположным суждениям. Если A знает о событии всё, что известно B и что-то ещё, его суждение о нём является более обоснованным, и именно его разумно принять. В обычной речи *шанс* и вероятность являются почти синонимами, и мы чаще всего употребляем их безразлично. Но если нужно различить их значение, мы здесь применяем слово *шанс* к самим событиям вне зависимости от наших знаний о нём, и оставляем за словом *вероятность* его предыдущее определение[1]. Итак, событие по своей природе обладает большим или меньшим шансом, известным или нет, а его вероятность зависит от наших знаний о нём.

Например, при игре в орлянку шанс появления возможных исходов зависит от строения подбрасываемой монеты, и можно полагать физически невозможным, что их шансы совпадают. Тем не менее, если строение монеты нам неизвестно, и мы ещё не испытали её, то для нас вероятности появления обоих исходов абсолютно равны. По существу, у нас нет никаких оснований верить в появление одного из них сильнее, чем в наступление другого, но положение меняется, если монета уже подбрасывалась много раз. Собственно шанс каждого исхода остался без изменения, однако для того, кто знает результат испытаний, вероятности этих исходов изменяются в зависимости от числа их появления.

2. Мерой вероятности события является отношение числа благоприятных для него случаев к общему числу всех равновозможных благоприятных и неблагоприятных случаев, обладающих равными шансами[2]. Это предложение означает, что, если для двух событий указанные отношения равны друг другу, то мы имеем одно и то же основание считать, что появится любое из них. В противном случае у нас окажется больше оснований считать, что наступит то из них, для которого это отношение больше.

Допустим, что в урне А находятся 4 белых и 6 чёрных шариков, а в урне В – 10 и 15 соответственно. Для извлечение белого шарика из урны А имеется 4 благоприятных случая из десяти, и 10 из 25 для появления такого же шарика из урны В. Отношения первого числа ко второму равны 2/5; прежде всего следует доказать, что вероятности извлечения белого шарика из обеих урн равны друг другу. [Следует полстраницы доказательства.]

Теперь я предполагаю, что в урне А находятся 4 белых шарика и 3 чёрных, а в урне В – 3 и 2 соответственно. […] У нас больше оснований […]. Дробь, соответствующая указанной мере вероятности, казалось бы, должна всегда быть соизмеримой [с единицей], но если для некоторого события числа благоприятных и всех случаев бесконечны, вероятность может оказаться несоизмеримой. Пусть s – протяжённость плоской поверхности, а σ – некоторый её кусок. Если бросать на s круглую монету так, что её центр может в равной степени оказаться в любой точке s, то очевидно, что вероятность падения монеты на некоторую точку σ будет равна σ/s при возможной несоизмеримости σ и s.

3. В приведенных примерах указывались определённые количества шариков, но легко видеть, что рассуждения были общими и не зависели от конкретных чисел. Можно также предположить, что рассматриваемое событие состоит в появлении белого шарика, а количества белых и чёрных шариков представляют благоприятные и неблагоприятных случаи. Чтобы упростить умозаключения, в каждой задаче о возможностях всегда можно принять подобную гипотезу по отношению к вещам совсем иной природы.

Пусть E – событие некоторого вида; обозначим через a число случаев, благоприятных его появлению, через b – число противоположных случаев и через p – вероятность E. Её мера (численное значение) равна $p = a/(a + b)$. И если F – противоположное событие и его вероятность равна q, то $p + q = 1$. Если у нас нет более сильных оснований полагать, что произойдёт E, а не F, то их вероятности равны […].

Вместо события, которое может наступить или нет, E может означать некоторую вещь, о которой следует узнать, истинна она или ложна. Тогда a и b окажутся количествами случаев, в которых мы полагаем её истинной или ложной, так что p и q выразят вероятности истинности и ложности E. Оценивая в каждом примере либо возможность, либо сомнение и критичность по числу благоприятных и неблагоприятных случаев для наступления E и F, и полагая достоверным, что этими числами являются a и b, мы получим шансы E и F, равные p и q. Если же второе условие не выполнено, то p и q окажутся лишь их вероятностями, возможно отличающимися, как мы разъяснили выше, от их неизвестных шансов. Неизменно требуется, чтобы все случаи, благоприятные и неблагоприятные, были равновозможными либо сами по себе, либо в соответствии с нашими знаниями.

4. В теории шансов [!] *достоверность* считается специальным случаем вероятности, при котором соответствующее событие не имеет против себя ни одного шанса. В исчислении достоверность представляется единицей, тогда как некоторая вероятность – дробью, меньшей единицы. Полное *недоумение* при выборе между двух вещей выражается числом 1/2, а *невозможность* – нулём[3].

Указанное понятие о достоверности для нас достаточно. У нас нет нужды определять её саму по себе, что и невозможно.

Абсолютная достоверность принадлежит к неопределяемым вещам [понятиям], и можно лишь привести её примеры. Имеется очень немного вещей, достоверных в строгом смысле, как наше собственное существование. Далее, имеются некоторые аксиомы, не только достоверные, но и очевидные, и предложения, подобные, например, геометрическим теоремам, истинность которых мы доказываем (или доказываем, что противоположное невозможно).

Вещи, не противоречащие общим законам природы, в которых мы удостоверяемся по многочисленным свидетельствам, и те, которые подтверждаются повседневным опытом, обладают, тем не менее, лишь очень высокими вероятностями, достаточными, однако, чтобы либо в обычной жизни, либо даже в физических и исторических науках не было необходимости отличать их от полной достоверности.

Исчисление вероятностей имеет целью определить в каждой проблеме возможности или сомнения отношение числа случаев, благоприятных для наступления некоторого события или истинности некоторой вещи, к числу всех возможных случаев[4]. Таким образом мы можем точно определить в соответствии с величиной этой дроби, более или менее приближающейся к единице[5], наше основание считать, что эта вещь истинна, либо ложна, что это событие произошло или наступит. И мы также сможем, нисколько не заблуждаясь, сравнивать это основание в двух полностью отличных друг от друга вопросах.

Исчисление вероятностей основано на небольшом числе правил, которые мы опишем и которые совершенно строго показывают, как следует рассматривать пример, предложенный в § 2. К этим принципам [правилам] следует относиться как к необходимому дополнению логики[6], потому что имеется такое большое число вопросов и методов рассуждения, которые не приводят к полной достоверности. Никакая другая часть математики не имеет такого большого числа приложений и не является столь же непосредственно полезной. Гл. 2 рассматривает отвлечённые и спорные проблемы общей философии и указывает их ясное и неоспоримое решение.

5. Если p и p_1 – вероятности независимых событий[7] E и E_1, вероятность их совместного появления, т. е. события, состоящего из них, равна pp_1. По существу, если событие E состоит в извлечении белого шарика из урны A, содержащей a белых и b чёрных шариков при $a + b = c$, а E_1 – в появлении такого же шарика из урны A_1, содержащей, соответственно, a_1 и b_1 шариков при $a_1 + b_1 = c_1$, то, по предыдущему,

$$p = a/c, p_1 = a_1/c_1.$$

Составное событие, – извлечение двух белых шариков, по одному из каждой урны. При извлечении шариков наугад каждый шарик из А может появиться с каждым из A_1, так что оказывается cc_1 равновозможных случаев. Из них число случаев, равное числу комбинаций белых шариков из обеих урн, окажется благоприятным составному событию. Поэтому (§ 2) оно имеет вероятность $aa_1/cc_1 = pp_1$. Мы также видим, что если $p, p_1, p_2, \ldots$ – вероятности некоторого числа независимых друг от друга событий $E, E_1, E_2, \ldots$ то вероятность их совместного появления, т. е. указанного составного события, равна $pp_1p_2 \ldots$ Это общее положение можно вывести и из частного случая события, составленного из двух других. Если pp_1 – вероятность совместного появления E и E_1, то совместное наступление его и E_2 будет иметь вероятность pp_1p_2 и т. д.

Все дроби $p, p_1, p_2, \ldots$ менее единицы, по крайней мере если ни одно из событий $E, E_1, E_2, \ldots$ не является достоверным. Следовательно, вероятность совместного события к тому же меньше вероятности каждого исходного. Она снижается всё более и более по мере возрастания числа исходных событий; вообще, если количество последних становится бесконечным, она стремится к нулю и оказывается в точности нулём или бесконечно малой величиной. Исключением является лишь случай, при котором бесконечный ряд вероятностей состоит из членов, безгранично приближающихся к единице, т. е. к достоверности. Их произведение может оказаться конечной величиной, меньшей единицы. Обозначим, к примеру, положительную величину меньшую или, самое большее, равную единице, через α, и пусть тогда

$$p = \alpha, p_1 = 1 - \alpha^2, p_2 = 1 - \alpha^2/4, p_3 = 1 - \alpha^2/9, \ldots,$$

и произведение этих вероятностей (или вероятность составного события) будет по известной формуле равно $(1/\pi)\sin\alpha\pi$. Здесь, как обычно, π – отношение окружности к своему диаметру.

6. Вот задача о вероятности составного события, которую мы решим в качестве примера на предыдущее правило. Я предполагаю, что вычитаю друг из друга соответствующие цифры двух случайно выбранных чисел. Требуется определить вероятность, что вся операция возможна без увеличения большего числа при каком-либо вычитании цифр друг из друга. Соответствующие большие и меньшие цифры могут принимать 10 значений от нуля до девяти. Поэтому при каждом вычитании имеется 100 различных равновозможных случаев. Требуется, чтобы большая цифра превышала меньшую или была ей равна, что происходит в 55 случаях из 100. Именно, в одном случае, когда большая цифра равна 0, в двух случаях когда она равна 1, …, в 10 случаях – 9. Количества этих случаев образуют арифметическую прогрессию из 10 членов, и её сумма равна $(1/2)\cdot10\cdot(1 + 10) = 55$. При каждом вычитании вероятность обойтись без увеличения большей цифры равна 0.55, а соответствующая вероятность для всех вычитаний числом i равна 0.55^i.

Пусть требуется вычитать друг из друга мантиссы логарифмов в таблице Callet[8]. Мы имеем

$$i = 7,\ 0.55^i = 0.0152243\ \ldots,$$

т. е. вероятность, заключённую между 1/66 и 1/65. Ту же величину 0.55^i мы получим для вероятности сложения двух чисел из i цифр каждое без того, чтобы при каком-либо сложении требовалось запоминать 1.

7. Если последовательно m раз появляется одно и то же событие E, то при постоянном p произведение $pp_1p_2\ \ldots$ становится равным p^m. Если E и F – противоположные события, вероятности которых p и q, то $p + q = 1$ (§ 3). Если эти шансы [?] остаются постоянными в течение $(m + n)$ испытаний, то произведение p^mq^n окажется вероятностью появления E и F соответственно m и n раз […].

Порядок следования событий E и F не влияет на эту вероятность, т. е. на вероятность составного события […]. Но если этот порядок не задан, то ясно, что вероятность соответственного составного события окажется выше, чем при любом определённом порядке и будет равна некоторому кратному числа p^mq^n; ниже мы укажем соответствующее общее выражение.

Если шансы E и F равны, то $p = q = 1/2$, то, принимая $m + n = \mu$, мы получим $(1/2)^\mu$ для вероятности появления в некотором порядке E и F m и n раз соответственно. Она не зависит ни от порядка наступлений этих событий, ни от соотношения количеств этих наступлений и изменяется только с полным числом испытаний μ. Такой случай происходит, если урна содержит равные количества белых и чёрных шариков при μ последовательных тиражах с возвратом. Вероятность появления μ белых шариков равна вероятности появления m белых и n чёрных в некотором порядке. Если μ – большое число, та и другая очень низки, но ни одна не ниже другой.

Перед началом тиражей у нас нет ни больше, ни меньше оснований считать появление ряда шариков одного и того же цвета более или менее вероятным, чем того же числа белых и чёрных шариков в произвольно установленном порядке. Однако, допустим, например, что извлечено подряд 30 шариков одного и того же цвета, и мы уверены, что количества шариков обоих цветов неизменно совпадают, или что появилось совсем иное и в каком-то смысле симметричное составное событие, как 30 шариков попеременно белого и чёрного цвета, или 15 белых, а затем 15 чёрных. Тогда мы чувствуем уверенность, что эти регулярные события не случайны, что тот, кто производил тиражи, знал цвет каждого из них и выбирал его в соответствии с какой-то целью. В подобных случаях вмешательство не случайной причины, оказывается, обладает вероятностью, очень близкой к достоверности, что мы покажем в дальнейшем[9].

8. Степень q^n есть вероятность наступления события F n раз подряд, без перерыва. Вычитая её из единицы, мы получим вероятность противоположного события, т. е. того, что в последовательности испытаний по меньшей мере один раз появится E. Следовательно, обозначив через r вероятность этого совместного события и полагая $q = 1 - p$, мы получим

$$r = 1 - (1 - p)^n.$$

Приравнивая эту вероятность половине, мы определим число испытаний, необходимых для того, чтобы иметь равное основание считать, что E появилось или нет. Иначе говоря, можно держать пари на равных, что E наступит по меньшей мере один раз. Итак,

$$(1 - p)^n = 1/2, \; n = - \lg 2/\lg(1 - p).$$

Если E – появление шестёрки или какого-либо другого числа очков при игре в шестигранные кости, то $p = 1/6$, $n = 3.8018 \ldots$ Можно с выгодой держать пари на то, что шестёрка появится хоть один раз в четырёх бросках, однако соглашаться на три броска вместо четырёх окажется невыгодным. При броске двух костей двойная шестёрка имеет вероятность $p = 1/36$, $n = 24.614$ … Выгодно держать пари на её появление в 25 бросках, но невыгодно соглашаться на 24[10].

Общее выражение для r показывает, что, каким бы слабым шанс p некоторого события Е ни был, лишь бы он не равнялся в точности нулю, всегда можно будет принять число испытаний n настолько большим, что вероятность наступления Е хотя бы один раз сколь угодно близко приблизится к достоверности. Действительно, как бы мало ни отличалась от единицы дробь $(1 - p)$, можно всегда принять достаточно большое n, чтобы $(1 - p)^n$ оказалось меньше заданной величины. В этом состоит существенное отличие абсолютно невозможной вещи от события Е, шанс p которого исключительно мал. Невозможная вещь не происходит никогда, а в достаточно длинном ряде испытаний событие со сколь угодно низкой вероятностью весьма вероятно наступит.

По биномиальной формуле

$$(1-p)^n = 1 - np + C_n^2 p^2 - C_n^3 p^3 + ...,$$

а если n – очень большое число, то $n - 1$, $n - 2$, … можно заменить на n, так что

$$(1-p)^n = 1 - np + C_n^2 p^2 - C_n^3 p^3 + ...$$

Это – ряд для e^{-np} (e – основание неперовых логарифмов) и таким образом приближённо

$$r = 1 - e^{-np}.$$

Если $p = 1/n$, это значение будет равно $(e - 1)/e$ и при очень больших n шанс события Е окажется равным $1/n$. Подобное n окажется достаточным для получения вероятности $(e - 1)/e$ почти равной 2/3 хотя бы для однократного появления Е.

9. Если события Е и E_1 никак не являются независимыми[11], т. е. если наступление одного из них влияет на шанс другого, вероятность совместного события Е и E_1 будет равна pp_1, где p – вероятность события Е, которое должно появиться первым, а p_1 – вероятность, что после этого наступит E_1.

Итак, если a и b – количества белых и чёрных шариков в урне А и $c = a + b$, и если Е – появление белого шарика при первом тираже, а E_1 – его появление при втором извлечении, перед которым извлечённый шарик не возвращается, то $p = a/c$, однако $p_1 = (a - 1)/(c - 1)$, так что

$$pp_1 = \frac{a(a-1)}{c(c-1)}$$

и аналогично

$$pp_1 = \frac{ab}{c(c-1)}$$

для извлечения белого и чёрного шариков в определённом порядке без возвращения первого из них.

Вообще, если произведено $m + n$ последовательных тиражей без возвращения и w – вероятность извлечения в произвольно установленном порядке m белых и n чёрных шариков, то

$$w = \frac{a(a-1)...(a-m+1)b(b-1)...(b-n+1)}{c(c-1)...(c-m-n+1)}.$$

Если в m_1+n_1 первых тиражах было извлечено m_1 белых и n_1 чёрных шариков, в числе $(c - m_1 - n_1)$ оставшихся окажется $a - m_1$ белых и $b - n_1$ чёрных. Вероятности извлечь белый и чёрный шарики теперь будут равны

$$\frac{a-m_1}{c-m_1-n_1}, \quad \frac{b-n_1}{c-m_1-n_1}.$$

Примем для m_1 последовательно все [целые] числа от 0 до $m - 1$ и от 0 до $n - 1$ для n_1, тогда произведение $(m + n)$ полученных таким образом величин, очевидно, составит w, что совпадёт с только что приведенной формулой. Если неизменно возвращать в урну извлекаемые шарики, их шансы будут оставаться постоянными и равными a/c и b/c, и вероятность извлечь m белых и n чёрных шариков в определённом порядке будет $a^m b^n/c^{m+n}$. Вот каким образом преобразуется выражение для w, если числа a и b очень велики и могут считаться бесконечными по сравнению с m и n, так что шансы белого и чёрного шарика в течение всех испытаний остаются неизменными. Полагая $n = 0$, мы получаем вероятность извлечения подряд m белых шариков

$$w = \frac{a(a-1)...(a-m+1)}{c(c-1)...(c-m+1)}.$$

Вместо урны A рассмотрим, к примеру, игру с 16 красными и столькими же чёрными картами. Требуется определить вероятность извлечения всех красных карт подряд. Здесь $a = 16$, $c = 32$, $m = 16$ и

$$w = \frac{16!}{17 \cdot 18 \cdot ... \cdot 31 \cdot 32} = \frac{1}{601080390}.$$

Эта величина немного менее $1/60 \cdot 10^7$. Поэтому требуется немного больше $60 \cdot 10^7$ испытаний, чтобы иметь вероятность 2/3 для осуществления указанного результата[12].

10. Пусть событие E может наступить многими различными и независимыми друг от друга способами с вероятностями p_1, p_2, … Полная вероятность p окажется равной их сумме. Конкретно предположим, что имеется i урн с белыми и чёрными шариками числом c_1 в первой урне, из которых a_1 белых, c_2 и a_2 во второй урне и т. д. Событием E назовём извлечение белого шарика из урны, выбранной из всех наугад с одной и той же вероятностью. Таким образом, оно может наступить i различными способами и в соответствии с правилом § 5 вероятности p_1, p_2, … окажутся равными

$$p_1 = \frac{a_1}{ic_1},\ p_2 = \frac{a_2}{ic_2},\ ...$$

Требуется доказать, что

$$p = \frac{1}{i}[\frac{a_1}{c_1} + \frac{a_2}{c_2} + ...].$$

Доказательство основано на лемме, которая будет столь же полезна и в других случаях. Пусть имеется i урн C с белыми и чёрными шариками в различных соотношениях, общим числом μ в каждой. Вероятность извлечь белый шарик из какой бы то ни было урны не изменится, если объединить $i\mu$ шариков в единой урне B. По существу они тогда образуют расположенные каким-то образом группы, каждая из которых содержит шарики числом μ, взятые из одной и той же исходной урны.

Этого достаточно, чтобы шанс извлечь шарик из любой группы был равен $1/i$, т. е. таким же, как в предшествующем случае, когда каждая группа находилась в одной из исходных урн. Шанс извлечь белый шарик из случайно выбранной группы не изменяется, а потому вероятность этого события для урны B та же, что и прежде. Это заключение уже не имеет места, если количества шариков в урнах различны. Шансы выбрать одну из этих урн одни и те же и равны $1/i$, но теперь, когда все шарики объединены в урне B, их группы не равны количественно, и шансы случайного выбора групп различаются друг от друга; чем многочисленнее группа, тем, очевидно, этот шанс для неё больше.

Итак, приведём все дроби a_1/c_1, a_2/c_2, … к общему знаменателю, который мы обозначим через μ, а их числители через α_i. Шанс извлечь белый шарик из каждой урны, а потому и из их множества, не изменится, если количества c_1, c_2, … белых и чёрных шариков заменить единым числом μ, а количества белых шариков a_1, a_2, … – числами α_1, α_2, … Вероятность извлечь белый шарик не изменится, если теперь объединить все шарики в одной урне C. Она будет содержать $i\mu$ шариков, включая $\alpha_1 + \alpha_2 + \ldots$ белых. Указанная вероятность окажется равной

$$\frac{1}{i}\left[\frac{\alpha_1}{\mu} + \frac{\alpha_2}{\mu} + \ldots\right]$$

и, в соответствии с предыдущими уравнениями, совпадёт с p, ч. т. д.

11. Чтобы применить это правило к примерам, предположим прежде всего, что кто-то знает, что шарик извлечён либо из урны A, содержащей 5 белых и 1 чёрный шарик, либо из урны B с тремя и четырьмя шариками соответственно, притом что у него нет никаких оснований считать, что шарик скорее извлекли из одной, а не из другой урны. Для него вероятность извлечения белого шарика равна

$$w = \frac{1}{2}\cdot\frac{5}{6} + \frac{1}{2}\cdot\frac{3}{7} = \frac{53}{84}.$$

Для другого лица, который знает, что шарик извлекли из урны B, вероятность извлечь чёрный шарик равна $p = 4/7 = 48/84$.

Каждая из этих вероятностей превышает 1/2, и потому первый из двух лиц полагает, что был извлечён белый шарик, второй же считает, что появился чёрный. Мы должны принять мнение второго лица, потому что он был более осведомлён, хоть вероятность 48/84 ниже, чем 53/84.

Это – очень простой пример, и легко можно было бы привести много подобных сказанному в § 1 о противоположных мнениях лиц, имеющих различное познание по одному и тому же вопросу. Допустим ещё, что мы знаем, что урна А содержит некоторое число n белых и чёрных шариков, соотношение которых нам совершенно неизвестно. Мы поэтому можем сформулировать о нём $n + 1$ различных и равновозможных гипотез: n белых шариков; $n - 1$ белых и 1 чёрный; $n - 2$ и 2 и т. д., n чёрных шариков. Вероятности каждой гипотезы равны $1/(n + 1)$ и частные вероятности извлечь белый шарик будут равны

$$p_1 = \frac{1}{n+1}\frac{n}{n},\ p_2 = \frac{1}{n+1}\frac{n-1}{n},\ \ldots,$$

а полная вероятность окажется равной

$$w = \frac{1}{n+1}\left[\frac{n}{n} + \frac{n-1}{n} + \ldots + \frac{n-n}{n}\right] = \frac{1}{2},$$

как и должно было быть, потому что у нас нет никаких оснований считать, что белый шарик появится скорее, чем чёрный.

Но если мы знаем, что белых шариков в урне А наверняка больше, чем чёрных, то $w > 1/2$. При вычислении этой вероятности следует различать нечётные и чётные n. Пусть i – некоторое целое число и $n = 2i + 1$. Можно будет сформулировать только $i + 1$ различных и равновозможных гипотез, а именно: $2i + 1$ белых шариков; $2i$ белых и 1 чёрный, … $i + 1$ белых и i чёрных. В этом первом случае полная вероятность

$$w = \frac{1}{(i+1)(2i+1)}[(2i+1) + 2i + (2i-1) + \ldots + (i+1)] = \frac{1}{2}\frac{3i+2}{2i+1}.$$

При $i = 0$ эта вероятность равна 1, как и должно было быть; она безгранично приближается к 3/4, неизменно убывая по мере возрастания i. В случае $n = 2i + 2$ можно сделать $i + 1$ равновозможных гипотез. Именно, можно предположить, что А содержит $2i + 2$ белых шариков; $2i + 1$ белых и 1 чёрный шарик; … $i + 2$ белых и i чёрных. В результате

$$w = \frac{1}{(i+1)(2i+2)}[(2i+2)+(2i+1)+2i+...+(i+2)] = \frac{1}{2}\frac{3i+4}{2i+2}.$$

Как и в предыдущем случае, $w = 1$ и 3/4 при крайних значениях $i = 0$ и ∞. Для любого иного целого i эта вероятность превышает предыдущую на

$$\frac{i}{4(i+1)(2i+1)}.$$

Максимум этой разности равен 1/24 при $i = 1$.

Урна А содержит c шариков, из которых a белых. Шарики разделены на группы, в которых находятся, соответственно, c_1 и a_1, c_2 и a_2, … шариков, причём

$$c_1 + c_2 + \ldots = c,\ a_1 + a_2 + \ldots = a.$$

Вероятность извлечь белый шарик из урны должна быть равна a/c, что просто подтверждает правило § 10. Он может быть извлечён из первой группы; шанс этого события равен произведению вероятности c_1/c выбора этой группы и шанса a_1/c извлечь из неё белый шарик. Для других групп рассуждение аналогично, и потому полная вероятность извлечения белого шарика равна

$$p = \frac{c_1}{c}\cdot\frac{a_1}{c_1} + \frac{c_2}{c}\cdot\frac{a_2}{c_2} + ... = \frac{a}{c}.$$

Но если поместить эти группы в различные урны A_1, A_2, … при неравных c_1, c_2, … шанс извлечь белый шарик не будет равен a/c. В общем, он зависит от распределения белых и чёрных шариков в этих урнах; не зная его, мы не можем вычислить этот шанс. Однако, при случайном выборе урны основание считать, что будет извлечён белый шарик, окажется, очевидно, тем же самым, что и при извлечении такого же шарика из единой урны А. Следовательно, вероятность извлечения белого шарика, отличная от шанса этого события, будет равна a/c. Пусть А содержит 2 белых шарика и 1 чёрный, из которых белые попадают в A_1, а чёрный – в A_2. Для несведущего лица существуют три равновозможных распределения шариков, а именно: 2 белых – в урне A_1 и чёрный – в A_2; 1 белый и 1 чёрный – в A_1 и 1 белый – в A_2; этот второй белый и чёрный – в A_1, и первый белый – в A_2. Вероятности извлечь белый шарик из одной из этих урн равна

$$\frac{1}{2}(1+0),\ \frac{1}{2}(\frac{1}{2}+1),\ \frac{1}{2}(\frac{1}{2}+1).$$

Одна треть этой суммы равна 2/3 – полной вероятности, как и в случае появления белого шарика из единой урны А.

Рассмотрим, наконец, систему урн D_1, D_2, …, в первой из которых находится c_1 шариков, a_1 из них – белые, во второй, соответственно, c_2 и a_2 и т. д. и предположим, что по какой-то причине урны выбираются из системы с неравными шансами. Пусть k_1 – вероятность выбора урны D_1, k_2 – выбора урны D_2 и т. д. По правилу § 5 вероятность извлечь белый шарик из первой урны будет равна $p_1 = k_1a_1/c_1$, из второй – $p_2 = k_2a_2/c_2$, и т. д. Полная вероятность этого события будет поэтому равна

$$w = k_1a_1/c_1 + k_2a_2/c_2 + \ldots$$

Рассмотрение системы урн A_1, A_2, …, для которой вероятности k_1, k_2, … окажутся равными, достаточно для доказательства правила § 10 во всей его общности. Это доказанное правило в применении к урнам D_1, D_2, …, для которых k_1, k_2, … принимают некоторые значения, приводят в общем случае, как сказано, к выражению для w в общем случае.

13. Пусть теперь события E и F взаимоисключающи и противоположны, одно из которых должно всегда наступать, а их вероятности пусть будут p и q, так что (§ 3) $p + q = 1$. Пусть каждое из этих событий может происходить различными способами с вероятностями p_1, p_2, … и q_1, q_2, … соответственно. Тогда, по предыдущему,

$$p = p_1 + p_2 + \ldots,\ q = q_1 + q_2 + \ldots,\ p_1 + p_2 + \ldots + q_1 + q_2 + \ldots = 1.$$

В различных проблемах о возможностях члены последнего уравнения являются вероятностями различных случаев, благоприятных и неблагоприятных для появления Е, и поэтому их сумма всегда должна быть равна единице или достоверности. Это должно иметь место, если учесть все возможные случаи. В соответствии с этим же уравнением выражение для p можно представить в форме

$$p = \frac{lp_1 + lp_2 + ...}{lp_1 + lp_2 + ... + lq_1 + lq_2 + ...},$$

где l может быть произвольным. Члены этой дроби пропорциональны шансам p_1, p_2, …, q_1, q_2, … Полученное выражение можно представить в виде

$$p = \frac{\alpha_1 a_1 + \alpha_2 a_2 + ...}{\gamma_1 c_1 + \gamma_2 c_2 + ...}.$$

Таким образом, если случаи, благоприятные и неблагоприятные событию E, имеют неравные шансы, то его вероятность можно вычислить, умножая число равновероятных случаев на количества, пропорциональные их соответствующим вероятностям и разделив сумму этих частных произведений на сумму тех же произведений для всех возможных случаев.

Это правило является более общим и часто более удобным, чем данное в § 2, в котором было предусмотрено равенство шансов всех благоприятных и неблагоприятных случаев.

14. Правила §§ 5 и 10 достаточны для получения формул, относящихся к повторениям в серии испытаний события с известными неизменными или изменяющимися шансами.

Будем как всегда обозначать противоположные события какой-либо природы через E и F, одно из которых должно наступать при каждом испытании. Прежде всего предположим, что их вероятности постоянны и известны, и пусть в каждом испытании их шансы будут p и q. Число испытаний, в которых указанные события произошли m и n раз соответственно, обозначим через μ. Тогда

$$p + q = 1, m + n = \mu.$$

Вероятность этого результата, происшедшего в определённом порядке, не зависит от конкретного порядка и равна $p^m q^n$ (§ 7). Поэтому, обозначив её через П и через K – количество способов [порядков], в которых может быть получен указанный результат, получим (§ 10)

$$\Pi = Kp^m q^n.$$

Чтобы определить K, я вначале предположу, что все μ события A, B, C, D, …, которые должны были произойти, различны. Тогда K окажется равным числу *перестановок* μ букв, т. е. $\mu!$. […] Если m событий A, B, C, D, … представляют собой одно и то же событие E, то соответствующее число перестановок будет равно $\mu!/m!$. Если F произошло $\mu - m = n$ раз, то для него это число равно $m!/n!$. Если же E произошло m раз, а F – n раз, то число перестановок (т. е. искомое значение K) равно

$$K = \frac{\mu!}{m!n!}.$$

Но K симметрично относительно m и n, а потому $K = C_\mu^m = C_\mu^n$. Таким образом, П есть $(m + 1)$-й член разложения $(p + q)^\mu$ по возрастающим степеням p или $(n + 1)$-й член – по возрастающим степеням q. Отсюда следует, что в нашем случае, когда шансы p и q событий E и F постоянны, шансы всех составных событий, которые могут наступить при μ испытаниях, выражаются различными членами формулы $(p + q)^\mu$. Число этих событий равно $\mu + 1$. Их вероятности различны, либо ввиду множества комбинаций, которые могут приводить к ним, и которые выражаются числом K, либо ввиду неравенства шансов p и q. При $p = q$ и чётном μ наиболее вероятное событие соответствует случаю $m = n$, и одному из двух чисел, соответствующих равенствам $m - n = \pm 1$, при нечётном μ.

15. Обозначим через P вероятность того, что E наступит не менее m раз в μ испытаниях. Это составное событие может произойти $m + 1$ различными способами, а именно когда E наступит μ, $\mu - 1$, … и, наконец, $\mu - n$ или m раз. Соответствующие вероятности определяются из предыдущего выражения для П при подстановке μ и 0, $\mu - 1$ и 1, $\mu - 2$ и 2, …, m и n. По правилу § 10 полное значение P будет равно сумме этих $n + 1$ частных вероятностей, т. е.

$$P = p^\mu + \mu p^{\mu-1}q + C_\mu^2 p^{\mu-2}q^2 + \ldots + C_\mu^n p^m q^n,$$

или иначе, сумме первых $n + 1$ членов разложения $(p + q)^\mu$, расположенного по возрастающим степеням q. Для $m = 0$ и $n = \mu$

$P = (p + q)^\mu = 1$, что и должно было быть, потому что вероятность составного события, включающего все сочетания E и F, которые могут произойти, должна обратиться в достоверность.

Событие $m = 1$ противоположно появлению F во всех испытаниях, и в этом случае P выражается всем разложением $(p + q)^\mu$ без последнего члена q^μ, что соответствует значению r в § 8. Если $\mu = 2i + 1$, то вероятность, что E наступает чаще, чем F, определяется по общему выражению для P при $m = i + 1$ и $n = i$. При $\mu = 2i$, если принять, что в этом же выражении [для P] $m = n = i$, получается вероятность, что E наступит по меньшей мере столько же раз, сколько F.

16. Из этой формулы можно также вывести решение первой исторически рассмотренной задачи на вероятность, которую мы указали в самом начале Предисловия, и которая известна под названием *раздела ставки*. Два игрока, A и B, играют в некоторую игру, в которой после каждой партии один из них должен выиграть очко, и соответствующие вероятности равны p и q. Чтобы выиграть игру в целом, игроку A осталось набрать a очков, а игроку B – b очков. Требуется определить вероятности α и β этих выигрышей. Одно из возможных противоположных событий должно произойти, так что $\alpha + \beta = 1$, и потому достаточно определить α.

Прежде всего заметим, что игра закончится не более, чем после $a + b - 1$ партий, потому что A выиграет не менее a, а B – не менее b. Более того, не меняя шансов этих выигрышей ни одним из них, игроки могут договориться играть именно столько партий, потому что только один из них может набрать нужное число очков. Если A набрал a очков до того, как B набрал b, или B набрал b очков до того, как A набрал a, либо A, либо B уже выиграли игру, что бы ни случилось в дальнейшем.

Поэтому, чтобы определить шансы α и β, мы можем предположить, что игроки всегда играют $a + b - 1$ партий. Тогда α – вероятность, что событие E, шанс которого равен p при каждом испытании, наступит не менее a раз, и она определяется по предыдущему выражению для P при $\mu = a + b - 1$, $m = a$, $n = b - 1$. Если, например, $p = 2/3$, $q = 1/3$, $a = 4$ и $b = 2$, то

$$\alpha = 112/243,\ \beta = 131/243,\ \beta > \alpha.$$

Игрок A, вдвое искуснее, чем B, т. е. имеющий вдвое больший шанс выиграть любую партию, тем не менее не может без убытка держать пари на выигрыш четырёх партий до того, как B выиграет две[13].

Если игроки договорились прекратить игру до её окончания, то, как будет сказано ниже, игроку A достанется *ставка*, умноженная на шанс α её выигрыша, а B – ставка, умноженная на шанс β. Иными словами, ставка делится в отношении $\alpha:\beta$.

17. Пусть вместо событий E и F имеется большее их число, например, три, E, F, G, только одно из которых должно наступать в каждом испытании. Их постоянные вероятности p, q, r и μ – число испытаний. По нетрудному обобщению метода § 14 вероятность их наступления m, n, o раз соответственно будет равна

$$\frac{\mu! p^m q^n r^o}{m! n! o!}, \; p + q + r = 1, m + n + o = \mu,$$

т. е. общему члену разложения $(p + q + r)^\mu$.

Таков случай урны, содержащей шарики трёх цветов в соотношениях, указанных дробями p, q, r, если событиями E, F, G считаются появления шариков соответствующих цветов с возвратом. В указанном разложении трёхчлена сумма членов, содержащих p в степени, равной или большей m, будет вероятностью того, что E наступит не менее m раз в μ испытаниях. При любом числе событий, только одно из которых наступает при каждом испытании, можно таким образом немедленно определить эту вероятность по предыдущему выражению для P. По существу, неизменно обозначая через $p, q, r, \ldots$ постоянные шансы появления в каждом испытании одного из событий E, F, G, …, допустимо считать это составным событием, которое я обозначу F′, а его вероятность – q'. Тогда[14]

$$q' = q + r + \ldots, \; p + q' = 1.$$

E и F′ – противоположные события, только одно из которых наступает в каждом испытании. Следовательно, вероятность появления E по меньшей мере m раз в μ испытаниях определяется подстановкой q' вместо q в выражение для P. Для примера на это правило, основанное на разложении степени многочлена, я предположу, что урна A содержит некоторое число m последовательно пронумерованных шариков $(1, 2, \ldots, m)$; μ раз подряд шарик извлекается с возвращением. В каждом тираже шанс появления шарика с определённым номером один и тот же для всех шариков, постоянен и равен $1/m$.

Пусть теперь

$$n_1, n_2, \ldots, n_m \tag{17.1}$$

данные равные или неравные числа 0, 1, 2, … и

$$n_1 + n_2 + \ldots + n_m = \mu.$$

Пусть, далее, U – вероятность, что в каком-то порядке шарик № 1 будет извлечён n_1 раз, шарик № 2 – n_2 раз, …, шарик № m – m раз, и пусть

$$(t_1 + t_2 + \ldots + t_m)^\mu = \theta.$$

Раскладывая θ по степеням и произведениям неопределённых величин $t_1, t_2, \ldots, t_m$, получим U как член этого разложения, содержащий произведение

$$t_1^{n_1} t_2^{n_2} \ldots t_m^{n_m},$$

если полагать, что каждая из этих величин равна $1/m$. Обозначив численный коэффициент этого произведения через N, получим

$$U = N/m^\mu.$$

N – целое число, зависящее от μ и чисел (17.1) и равное

$$N = \frac{\mu!}{n_1!\, n_2! \ldots n_m!},$$

причём во всех подобных произведениях мы считаем, что $0! = 1$.

Пусть теперь s – сумма номеров, извлечённых при μ тиражах

$$s = n_1 + 2n_2 + \ldots + mn_m.$$

Если эта сумма задана, а для чисел (17.1) приняты последовательно все целые числа и нуль, удовлетворяющие этому уравнению и в сумме равные μ [что было обусловлено выше], и через N', N'', … обозначены соответствующие значения N, а через V – сумма соответствующих U, тогда

$$V = (1/m^{\mu})(N' + N'' + \dots)$$

окажется вероятностью получить в μ тиражах заданную сумму s.

Вычислить V проще при замене в выражении для θ неопределённых величин t_1, t_2, …, t_m на степени t, t^2, …, t^m. Обозначив теперь через Т новое значение θ, получим[15]

$$Т = (t + t^2 + \dots, + t^m)^{\mu}.$$

Легко видеть, что сумма $N' + N'' + \dots$ оказывается числовым коэффициентом M при t^s в разложении Т, и поэтому

$$V = M_s/m^{\mu}.$$

Этот коэффициент зависит от данных чисел μ, m, s и в каждом примере его легко вычислить. Вместо одной-единственной урны можно предположить μ урн A_1, A_2, …, A_{μ}, каждая из которых содержит m шариков, пронумерованных подряд 1, 2, …, m, причём из каждой одновременно извлекается по одному шарику. Можно также заменить эти урны тем же числом костей, шесть граней каждой из которых пронумерованы от 1 до 6. Тогда $m = 6$, а V окажется вероятностью получить s очков при броске μ костей.

Пусть μ = 3, тогда

$$Т = t^3(1 + t + t^2 + \dots, + t^5)^3,\ V = M_s/6^3.$$

Разложение Т состоит из 16 членов; коэффициенты членов, равноудалённых от его концов, как M_3 и M_{18}, M_4 и M_{17}, …, M_{10} и M_{11}, попарно равны, а сумма всех коэффициентов равна значению Т при $t = 1$, т. е. 6^3. Суммы первых восьми, M_3, M_4 …, M_{10}, и последних восьми, M_{11}, M_{12}, …, M_{18}, равны $(1/2)6^3$. Поэтому при броске трёх костей вероятность набрать 10 или меньше очков равна 1/2, равно как и 11 или больше, и можно держать пари на равных, что сумма полученных очков будет либо более, либо менее десяти, и на этом факте основана игра passe-dix. Без всякого исчисления легко удостовериться в равенстве шансов обоих игроков, если заметить, что сумма очков каждой пары противоположных граней одной и той же кости равна семи (1 и 6, 2 и 5, 3 и 4)[16]. При броске трёх костей сумма их трёх старших и трёх младших номеров всегда равна 21 и поэтому, если первая сумма превышает 10, вторая окажется менее 10 и наоборот. […]

Чтобы установить шансы различных значений s от трёх до 18, необходимо вернуться к разложению T. Мы тогда получим

$$M_3 = M_{18} = 1,\ M_4 = M_{17} = 3,\ M_5 = M_{16} = 6,\ M_6 = M_{15} = 10,$$
$$M_7 = M_{14} = 15,\ M_8 = M_{13} = 21,\ M_9 = M_{12} = 25,\ M_{10} = M_{11} = 27.$$

Разделив эти числа на $6^3 = 216$, мы установим шансы всех этих сумм.

18. Если шансы события E изменяются в течение испытаний, вероятность его повторения некоторое число раз будет зависеть от закона этого изменения. Пусть, как в § 9, E – безвозвратное извлечение белого шарика из урны A с a белыми и b чёрными шариками, μ – число испытаний и w – вероятность появления m белых и n чёрных шариков в какой-то определённой последовательности. Значение w даётся формулой того же параграфа; оно не зависит от выбранной последовательности. Вероятность какой-то из них мы обозначим через $\Pi = Kw$, где K – то же, что и в § 14. Примём как обычно

$$m + n = \mu,\ a + b = c,\ a - m = a',$$
$$b - n = b',\ c - \mu = a' + b' = c'.$$

Тогда a', b', c' – количества оставшихся шариков и их сумма, заменяющие прежние значения a, b и c. Учитывая выражения для K и w, мы выпишем

$$\Pi = \frac{\mu!a!b!c'!}{m!n!a'!b'!c!},$$

что может быть легко распространено на случай шариков трёх и более различных цветов. Сокращая общие множители в числителе и знаменателе этой формулы, мы получим более простое выражение[17]

$$\Pi = C_\mu^n \frac{a(a-1)...(a-m+1)b(b-1)...(b-n+1)}{c(c-1)...(c-\mu+1)}.$$

Вероятность, что после μ тиражей будет извлечено не менее m белых шариков, окажется суммой $n + 1$ значений Π, соответствующих замене m и n на μ и 0; на $\mu - 1$ и 1; …, $\mu - n$ и n. Обозначив эту вероятность через P, мы получим

$$Pc(c-1)...(c-\mu+1) = a(a-1)...(a-\mu+1) +$$
$$\mu ba(a-1)...(a-\mu+2) + C_\mu^2 b(b-1)a(a-1)...(a-\mu+3) +$$
$$C_\mu^3 b(b-1)(b-2)a(a-1)...(a-\mu+4) +$$
$$... + C_\mu^n b(b-1)...(b-n+1)a(a-1)...(a-m+1).$$

В частном случае $m = 0$ и $n = \mu$ соответствующее $P = 1$ и поэтому

$$c(c-1)...(c-\mu+1) = a(a-1)...(a-\mu+1) +$$
$$\mu ba(a-1)...(a-\mu+2) + C_\mu^2 b(b-1)a(a-1)...(a-\mu+3) +$$
$$C_\mu^3 b(b-1)(b-2)a(a-1)...(a-\mu+4) + \ldots +$$
$$\mu b(b-1) \ldots (b-\mu+2a) + b(b-1) \ldots (b-\mu+1),$$

что совпадает с известной формулой, аналогичной биномиальной. В этой и во всех подобных формулах каждая величина типа $a(a-1)(a-1)(a-2) \ldots (a-m+1)$ есть произведение m сомножителей, которое должно быть принято равным единице при $m = 0$. Поэтому указанная формула не относится к случаю $\mu = 0$, и такое же исключение имеет место для бинома в нулевой степени.

19. Вместо μ последовательных тиражей без возврата можно без изменения соответствующей вероятности сразу же извлечь $m + n = \mu$ шариков, что подтверждается следующим образом[18]. Я называю через G_μ число групп, состоящих из μ шариков каждая, которые могут быть составлены из исходных c шариков: $G_\mu = C_c^\mu$. Можно сформировать эти группы из групп по $\mu - 1$ шариков, сочетая каждый из них с $c - \mu + 1$ остальными шариками, и образовать таким образом $(c - \mu + 1)G_{\mu-1}$ требуемых групп. При этом, однако, μ групп по $\mu - 1$ шариков каждая составят одну и туже группу из μ шариков, и поэтому следует разделить указанное произведение на μ:

$$G_\mu = \frac{c-\mu+1}{\mu} G_{\mu-1}.$$

Для $\mu = 1$ очевидно $G_1 = c$. Последовательно получим

$$G_2 = [(c-1)/2]G_1 = C_c^2,\ G_3 = [(c-2)/3]G_2 = C_c^3,$$
$$G_4 = [(c-3)/4]G_3 = C_c^4,\ \ldots$$

и, наконец, G_μ.

Обозначим теперь буквами G'_m и G''_n значения G_m при замене *c* и μ соответственно на *a* и *m* и на *b* и *n*. Тогда

$$G'_m = C_a^m,\ G''_n = C_b^n.$$

Произведение $G'_m G''_n$ будет числом групп из $(m + n) = \mu$ шариков, которые могут быть составлены из $a + b = c$ шариков в урне и которые состоят из *m* белых шариков и *n* чёрных. Вероятность П выбора одной из них при извлечении μ шариков одновременно равна их числу, делённому на число всех групп по μ шариков в урне, т. е.

$$\Pi = \frac{G'_m G''_n}{G_\mu},$$

что совпадает со значением П в § 18. Выражение для *P* в том же параграфе также является вероятностью получить не менее *m* белых шариков при одновременном извлечении μ из них.

20. В примере § 18 шанс события E изменялся во время испытаний, потому что при каждом новом тираже он зависел от числа предшествовавших появлений событий E и F. Существуют, однако, задачи, в которых шансы этих двух событий некоторой природы не зависят от предыдущих событий, но изменяются от одного испытания к другому. Пусть в серии μ произведенных или будущих испытаний p_1 и q_1 будут шансами E и F в первом испытании, p_2 и q_2, во втором, …, p_μ и q_μ, в последнем,

$$p_1 + q_1 = p_2 + q_2 = \ldots = p_\mu + q_\mu = 1.$$

Определим вероятность, что E наступит или наступило *m* раз, а F – (μ – *m*) раз в каком-либо порядке; пусть произведение *m* дробей $p_1 p_2 \ldots p_\mu$ будет называться P_m, а произведение *n* дробей $q_1 q_2 \ldots q_\mu$ – Q_n. Если P_m включало некоторое p_i, то q_i не было включено в Q_n, а если P_m не включало некоторое p_i, то q_i было включено в Q_n. Я умножаю P_m на Q_n и образую сумму всех возможных величин P_mQ_n, число которых в § 14 было обозначено через K. Эта сумма выражает требуемую вероятность.

Указанное правило можно высказать иначе, что окажется полезным в дальнейшем. Пусть u и v – неопределённые величины и

$$R = (up_1 + vq_1)(up_2 + vq_2) \dots (up_\mu + vq_\mu) -$$

произведение $\mu = m + n$ сомножителей[19]. Выполнив умножения, мы получили бы многочлен из $\mu + 1$ членов, расположенных по степеням u и v. Коэффициент при $u^m v^n$ в этом многочлене будет искомой вероятностью появления в некотором порядке E и F соответственно m и n раз. Пусть $m = 3$. Тогда

$$R = u^3 p_1p_2p_3 + u^2 v\,(p_1p_2q_3 + p_1p_3q_2 + p_2p_3q_1 + uv^2\,(p_1q_2q_3 + p_2q_1q_3 + p_3q_1q_2) + v^3 q_1q_2q_3.$$

Коэффициент при u^3 очевидно будет вероятностью трёхкратного появления E; коэффициент при u^2v – двукратного наступления E и одного появления F. Действительно, E может наступить в двух первых испытаниях, а F – в третьем; F – во втором, а E – в двух других; и F – в первом, а E – в двух остальных. Коэффициент при uv^2 выразит вероятность двукратного появления F и одного наступления E, и, наконец, коэффициент при v^3 очевидно равен вероятности трёхкратного появления F.

Если в каждом испытании событие E может появиться многими равновозможными способами, его шанс в соответствии с правилом § 10 будет равен сумме соответствующих вероятностей, делённой на их число. Вычитая этот средний шанс E из единицы, мы получим средний шанс F. Указанные средние шансы в каждом испытании позволяют вычислить вероятность появления E и F соответственно m и n раз в $m + n$ испытаниях, а также и наступление события, составленного из E и F. Как бы эти частичные шансы не изменялись от испытания к испытанию по числу [равновозможных шансов] и величине, если их средние шансы остаются постоянными, то вероятности составных событий будут следовать тем же законам, что и при постоянных шансах.

21. Одно из самых частых приложений исчисления вероятностей имеет целью определение пользы или вреда, связанных с возможными вещами в соответствии с выгодой или убытком, к которым они приводят, и шансами их наступления. Это основано на следующем правиле.

Пусть одно из некоторого числа событий E, F, G, H, … должно наступить. Их вероятности, сумма которых равна единице, обозначим через p, q, r, s, … Предположим также, что для одного человека выгода g связана с появлением E, с наступлением F – для другого и т. д. Если все эти лица заранее договорятся разделить g, или если они обязаны почему-либо так сделать, то выгоду следует разделить пропорционально вероятностям её получения: gp достанется первому, gq – второму и т. д.

Если m – число всех равновозможных случаев, среди которых a, b, c, d, … благоприятны соответственно событиям E, F, G, H, … то

$$a + b + c + d + \ldots = m, \ p = a/m,$$
$$q = b/m, \ r = c/m, \ s = d/m, \ \ldots$$

Если же число лиц равно m, и каждый из них выигрывает при наступлении одного из m возможных случаев, то, очевидно, g следует разделить между ними поровну, а в противном случае они получат pg, qg, … Это правило укажет, что именно достанется каждому игроку в соответствии с его вероятностью выигрыша, если игроки договорятся разойтись, не оканчивая игры. Кроме того, до начала игры ставки каждого игрока должны быть пропорциональны шансу его выигрыша, ибо, если вместо игры игроки договорятся разойтись, каждый должен получить свою ставку. В соответствии с предыдущим правилом возвращаемая сумма должна быть также равна сумме ставок, умноженных на вероятность выигрыша. Такие вероятности в чисто случайных играх зависят от принятых правил, и, если они не слишком сложны, могут быть вычислены заранее. Но если успех зависит от умения каждого игрока, вероятность выигрыша обычно основывается на репутации и может быть установлена с некоторой точностью лишь по длительному опыту.

Пусть E и F – противоположные события с вероятностями p и q. Если A ставит α на появление E, а B ставит β на наступление F, то для справедливости пари должно выполняться равенство $p\beta = q\alpha$. Но нельзя забывать, что, вообще говоря, указанные вероятности отличаются от соответствующих шансов. Если они основаны на одних и тех же познаниях у A и B, то пари справедливо, пусть бы оно даже оказалось благоприятным одному за счёт другого. В противном случае отношение $\alpha:\beta$ уже не будет соответствовать указанным вероятностям, и справедливое пари установить невозможно.

22. По формулам § 19 нетрудно определить различные существовавшие шансы в лотерее Франции, к счастью упразднённой недавним законом. Сравнивая их с отношениями выигрышей к ставкам, замечаешь, что эти отношения были намного меньше, чем должно было бы быть при справедливой игре. За счёт игроков лотерея получала непомерную и наказуемую законом выгоду.

Пусть n – число номеров в лотерее, m из которых выходят в каждом тираже, l – число выигрышных номеров, выбранных игроками, и λ – вероятность, что эти последние номера выйдут в тираж [и действительно выиграют]. В соответствии с формулами § 19 количества групп по l номеров, выбранных из n или из m равны C_n^l и C_m^l и

$$\lambda = C_m^l \div C_n^l = \frac{m(m-1)...(m-l+1)}{n(n-1)...(n-l+1)}.$$

Пусть ставка игрока равна единице, тогда ставка лотереи должна быть равна $(1 - \lambda)/\lambda$. В случае выигрыша лотерея должна, кроме того, возместить игроку его ставку, так что выплата должна быть равна

$$\mu = [(1 - \lambda)/\lambda] + 1 = 1/\lambda.$$

Пусть также x – число тиражей, необходимых для справедливого пари о том, что номера, выбранные игроком, выиграют хоть один раз. Тогда по правилу § 8, если λ очень малая дробь,

$$(1 - \lambda)^x = 1/2,\ x = \ln 2/\lambda = 0.69315/\lambda,$$

В лотерее Франции $n = 90$, $m = 5$. При игре на три выигрышных номера $l = 3$, так что

$$\lambda = 5{\cdot}4{\cdot}3/90{\cdot}89{\cdot}88,\ \mu = 11\ 748,\ x = 8143.13\ \ldots$$

При справедливой игре лотерея должна была бы выплачивать выигравшему игроку в 11 748 раз больше его ставки, фактически же он получал в 5500 раз больше (меньше половины указанного). Несоответствие было ещё значительнее при игре на 4 или 5 номеров и меньше, если игра шла на два номера или на один. Можно было с выгодой держать пари, что три номера выиграют хотя бы один раз в 8144 тиражах, но было убыточно рассчитывать на 8143 тиража. По поводу одного номера, выбранного заранее, мы имели бы [поскольку $n/m = 18$]

$$(1 - 1/18)^x = 1/2,\ x = \lg2/(\lg18 - \lg17) = 12\ 137\ \ldots$$

Было бы поэтому убыточно держать пари на равных, что этот номер выйдет в 12 тиражах, и выгодно, если ставить на 13. Можно было бы также держать пари на равных, что все 90 номеров выйдут хоть раз в 85 или 84 тиражах (Laplace 1812/1886, с. 200 – 201)[19].

Некоторые игроки предпочитают выбирать номер, который долгое время не выигрывал, другие же, напротив, останавливаются на тех, которые выходили в тираж чаще остальных. Эти системы ошибочны в одинаковой мере[20]. Пусть, например, некоторый номер выиграет хоть один раз в 100 последовательных тиражах с вероятностью, весьма близкой достоверности и равной $1 - (17/18)^{100} \approx 0.997$. Если же он не выиграл при первых 88 извлечениях, вероятность его выхода в 12 последующих тиражах, как и для любого другого определённого номера, неизменно окажется равной почти 1/2. Что же касается наиболее часто выходящих номеров, это обстоятельство нельзя считать действием случая, сравнимого с очевидным равенством шансов всех номеров в каждом тираже. Ни в каких азартных играх, в которых равные или неравные шансы определённо известны, прошедшие события никак не влияют на вероятность будущих событий[21]. Никакие комбинации, воображаемые игроками, не могут ни увеличить выгоду, ни уменьшить потерю, которые происходят в соответствии с правилами предыдущего параграфа.

В общедоступных играх[22] в Париже выгода банкомёта в каждой партии незначительна. Так, в игре *Тридцать и сорок* она немного меньше 0.011 ставки, см. нашу статью (1825). Но ввиду быстроты игры и их большого числа, которые заканчиваются в течение нескольких часов, банкир получает верную прибыль, почти постоянную из года в год. Он может поэтому ежегодно уплачивать общественной администрации пять – шесть миллионов за предоставление ему монополии. Эти игры ещё вреднее, чем была лотерея, потому что деньги, которые проходят в них в одном только Париже, ежегодно достигают многих сотен миллионов и намного превосходят затрачиваемые в Лотерее по всей Франции.

Здесь не место обсуждать рассуждения, которые обычно приводятся в пользу сохранения общедоступных игр, и которые я никогда не мог считать убедительными. Достаточно сказать, что эти игры являются причиной многих несчастий и быть может преступлений, так что администрации следует их запретить, а не делить приносимую ими прибыль с теми, которым она продаёт исключительное право на них[23].

23. Произведение выгоды на вероятность её получения называется *математическим ожиданием*[24] каждого, кто заинтересован в некоторой рискованной операции. Пусть выгода составляет 60 000 франков, и 1/3 – вероятность события, с которым она связана. Тогда тот, кто должен будет при случае получить её, может считать её треть своей собственностью и должен включить эту треть в опись своего фактического состояния.

Вообще, если некто должен получить суммы g, g_1, … за появления событий E, E_1, …, шансы которых равны p, p_1, …, его ожидание будет равно $gp + g_1p_1 + …$ Если же одна или несколько величин g, g_1, … выражают потери, которых это лицо может опасаться, они войдут в указанную сумму со знаком минус, остальные же величины останутся положительными. В соответствии с величиной полного значения ожидания оно выражает прибавление или убыль состояния рискующего и, если он не желает ожидать исхода событий, должно считаться наравне с выставленными требованиями или долгами. Если выгоды или потери произойдут лишь после длительного времени, то, разумеется, их следует *дисконтировать* и тем самым вычислить их действительную стоимость, не зависящую от случайности.

Пусть g, g_1, … будет уплачено тому, кто оценивает своё состояние, лишь через n, n_1, … лет, тогда в данный момент эти выгоды должны быть разделены на $(1 + \theta)^n$, $(1+\theta)^{n_1}, \ldots$, где θ – годичная процентная ставка. Поэтому, обозначив через ε часть состояния этого лица, образованную соответствующими ожиданиями, мы получим

$$\varepsilon = \frac{gp}{(1+\theta)^n} + \frac{g_1p_1}{(1+\theta)^{n_1}} + \ldots$$

При учёте выгод и потерь следует считать ε суммой, которую кто-либо иной должен будет в этот момент либо уплатить рискующему, либо получить от него в зависимости от знака ε. Вычисление пожизненных рент на одного человека или на нескольких лиц и страхования жизни, а также установление пенсий основано на этой формуле и на *таблицах смертности*, как можно заметить в сочинениях, специально рассматривающих эти вопросы.

24. Поскольку польза от выгоды, происходящая для получателя, зависит от его состояния, относительную величину этой пользы, названную *моральным ожиданием*, отличают от математического ожидания. Если она бесконечно мала, то её отношение к действительному состоянию получателя является мерой морального ожидания, положительного или отрицательного. Применяя интегральное исчисление, можно вычислить эту меру, которая, оказывается, соответствует правилам благоразумного ведения рискованных операций.

Это исчисление также устанавливает причины отказа даже от справедливых игр, однако, видимо существуют и более весомые основания. Игры не создают ценностей, и это является неотразимым доводом против них, если они перестают быть просто развлечением. Более того, выгода выигрывавшего неизменно связана с несчастьем и быть может разорением проигравшего.

Торговля – это тоже игра, поскольку успех самых благоразумных рискованных операций всегда лишь очень вероятен и оставляет шанс для потери; умение и осмотрительность могут лишь ослабить его. Но торговля увеличивает стоимость вещей, их ведь перевозят с одного места на другое, и именно за счёт этого увеличения коммерсант получает свою выгоду и к тому же приносит пользу потребителю[25].

25. Правило § 21, хоть оно и простое и естественное, осложнено затруднением, которым иногда приходится немало заниматься. Пусть A и B играют в орлянку, приняв следующие условия. *Первое*, игра заканчивается, как только выпадёт орёл. *Второе*, B уплачивает A 2 франка, если это произойдёт при первом же броске, 4 франка, если при втором, …, и 2^n, если при n-м броске. *Третье*, игра заканчивается без всяких выплат, если орёл не выпадет в m первых бросках; без этого ограничения она могла бы длиться неопределённо долго.

Предполагается, что монета честная и шансы обоих игроков одни и те же[26]. Таким образом, вероятность орлу впервые наступить при n-м броске равна $1/2^n$. Это окажется возможным, если в $(n - 1)$ бросках появлялась решётка, вероятность чего равна $1/2^{n-1}$. После этого с вероятностью 1/2 появляется орёл, и $1/2^{n-1}\cdot 1/2 = 1/2^n$. В этом случае А получит 2^n франков, что добавит 1 франк к его математическому ожиданию. То же произойдёт в каждом из m возможных бросков монеты. Таким образом, полное математическое ожидание игрока А равно $1m$ франков, и для справедливости игры он должен дать В m франков, т. е. тысячу, миллион или бесконечную сумму, если игра могла бы продолжаться бесконечно. И тем не менее никто не согласился бы отдать какую-нибудь заметную сумму, например, тысячу франков, за игру в роли А, а это, видимо, указывает на недостаток правила математического ожидания. Чтобы устранить появившееся затруднение, были придуманы правило морального ожидания и его мера. Но следует заметить, что это затруднение основано на том, что В способен уплатить любую сумму, которую шансы игры доставят А. Каким бы большим ни было состояние В, оно ограничено. Пусть оно равно b, тогда А никогда не сможет получить больше, чем b, что весьма значительно уменьшает его математическое ожидание[27].

По существу, всегда

$$b = 2^{\beta}(1 + h),$$

где β – целое число, а h – положительное число, меньшее единицы. Если $\beta \geq m$, В сможет уплатить всё, что причитается А, но в противном случае, если орёл впервые появится после β первых бросков, это окажется невозможным. Математическое ожидание А в этих первых бросках равно β, после чего в $m - \beta$ последующих бросках оно оказывается равным b, умноженному на соответствующие вероятности от $1/2^{\beta+1}$ до $1/2^m$. Обозначим полное математическое ожидание А, т. е. то, что он должен дать В для справедливости игры, через ε, тогда

$$\varepsilon = \beta + \frac{1}{2}(1+h)[1+\frac{1}{2}+\frac{1}{4}+...+\frac{1}{2^{m-\beta-1}}] \;=\; \beta + (1+h)[1-\frac{1}{2^{m-\beta}}]$$

и не возрастает с m; напротив [?], эта величина почти не зависит от m и при очень больших его значениях примерно равна

$\varepsilon = \beta + 1 + h.$

Состояние B никогда не будет достаточным, чтобы β оказалось значительным, и A должен будет отдавать лишь небольшую сумму, заключённую между β + 1 и β + 2. Пусть B – банкир, обладающий сотней миллионов франков. Тогда β = 26 и игрок A окажется в невыгодном положении, если заплатит за игру 28 франков или более. Правило морального ожидания, приложенное к этой задаче, приводит к другому результату (Laplace 1812/1886, с. 451), который зависит от состояния A, а не B. Мне, однако, представляется, что сумма, которую должен уплатить A, ограничивается возможностями B.

26. Я заканчиваю эту главу несколькими замечаниями о влиянии неизвестного благоприятного шанса некоторого события, который неизменно повышает, как будет видно, вероятность повторения событий в серии испытаний[28]. К примеру, при игре в орлянку всегда можно считать, что ввиду физического состояния монеты какой-то из двух возможных исходов броска более вероятен. Но в любом случае вероятность появления одних и тех же исходов подряд повышается.

Назовём шансы возможных исходов $(1 \pm \delta)/2$, где δ – небольшая положительная дробь, значение которой неизвестно, неизвестно и какому именно исходу соответствуют эти шансы. Если совершается только один бросок, у нас не будет никаких оснований полагать, что исход, выбранный одним из игроков, либо более, либо менее вероятен, чем другой. Вероятность каждого равна 1/2 и δ как бы равно нулю. Но если требуется подбросить монету дважды, выгодно держать пари на повторение исходов, которое возможно в двух случаях из четырёх. Шансы повторения исходов равны $[(1 + \delta)/2]^2$ и $[(1 - \delta)/2]^2$ и вероятность одного из них по правилу § 10 равна сумме этих квадратов, т. е. $(1 + \delta^2)/2$. Другие два исхода имеют равные шансы $[(1 + \delta)/2][1 - \delta)/2]$ и вероятность этих несовпадающих исходов $(1 - \delta^2)/2$. Она ниже первой в соотношении

$$\frac{1-\delta^2}{1+\delta^2} = 1 - \frac{2\delta^2}{1+\delta^2}.$$

Если A ставит франк на совпадение исходов, а B – против него, то для справедливости пари ставка последнего должна быть соответственно меньше. При δ = 1/10 она примерно равна 98 сантимам. Если монета подбрасывается трижды, возможны 8 различных сочетаний, в том числе все три раза каждый из возможных исходов, остальные же 6 сочетаний включают несовпадающие исходы. При строгом равенстве $\delta = 0$ шансы всех восьми сочетаний совпадают, и при прежнем пари ставка A должна составлять треть ставки B. Но без сомнения $\delta \neq 0$ и соотношение ставок будет ещё благоприятнее для A, чем при двукратном испытании. Вероятности совпадения и несовпадения исходов равны

$$[\frac{1}{2}(1+\delta)]^3 + \frac{1}{2}(1-\delta)]^3 = \frac{1}{4}(1+3\delta^2);\;\; 1-\frac{1}{4}(1+3\delta^2) = \frac{3}{4}(1-\delta^2)$$

[…]. Это рассуждение легко обобщить на большее число испытаний, и, если угодно, на другие игры с тремя и более возможными исходами, неизвестные шансы которых могут не совпадать. Пусть умение в игре двоих имеет какое-то значение, притом равное умение обоих неправдоподобно. Не зная, кто из игроков более умел, можно держать пари на то, что один из них выиграет первые две партии. В противном случае держать такое же пари на более умелого не будет неизменно благоприятным. Действительно, из четырёх возможных сочетаний исходов только одно является таковым. Оно более вероятно, но возможно не сравнится с тремя другими вместе взятыми.

Вообще, пусть p – известная вероятность события E некоторой природы, а q – вероятность противоположного события F, $p + q = 1$. Предположим также, что существует причина, которая может повысить шанс какого-то одного события и понизить шанс другого на неизвестную дробь α. Пусть соответствующая вероятность того, что в m испытаниях могут появляться только указанные события, равна w, и повышен шанс наступления E. Тогда вероятность совпадения исходов m последовательных испытаний будет по правилу § 10 равна

$$(p + \alpha)^m + (q - \alpha)^m,$$

ибо и E, и F могут появляться в каждом из них. Если, напротив, повышен шанс появления F, соответствующая вероятность будет равна

$$(p - \alpha)^m + (q + \alpha)^m.$$

Не зная, шанс какого события повышен или понижен, мы полагаем, что эти два различных значения равновозможны и что поэтому вероятность каждого из них равна 1/2. И снова по правилу § 10 полусумма этих значений будет полной вероятностью совпадений: $w = P + Q$, или

$$w = (1/2)[(p + \alpha)^m + (q - \alpha)^m + (p - \alpha)^m + (q + \alpha)^m],$$
$$P = p^m + C_m^2 p^{m-2}\alpha^2 + C_m^4 p^{m-4}\alpha^4 + ...,$$
$$Q = q^m + C_m^2 q^{m-2}\alpha^2 + C_m^4 q^{m-4}\alpha^4 + ...$$

Если же $\alpha = 0$, то вероятность совпадений будет равна просто $p^m + q^m$. Таким образом, поскольку очевидно, что $w > p^m + q^m$, любая причина, которая повышает шанс какого-то одного из двух противоположных событий, повышает вероятность их совпадения в серии испытаний.

Примечания

1. Пуассон неукоснительно применял оба термина, *шанс* и *вероятность*, а в письме Курно 1836 г. указал, что *решительно настаивает* на отличии между ними (Курно 1843/1970, с. 29). Позднейшие авторы, однако, приняли гораздо более подходящие термины, *теоретическая* и *статистическая вероятности*.

2. Указанное ограничение автор повторил в конце § 3. Более общее определение, которое было известно уже Муавру (De Moivre 1712/1984, с. 237), см. § 13.

3. В теории информации половинная вероятность приравнивается незнанию.

4. По Чебышеву (1845/1951, с. 29), теория вероятностей имеет целью определение вероятности события по его связи с другими событиями с известными вероятностями. У Лапласа подобного определения не было.

5. Почему *приближается к единице*? Впрочем, практически важны именно подобные случаи; в косвенной форме это указал, например, Гнеденко (1950/1954, с. 184).

6. Это утверждение следует отметить.

7. Определение независимых событий, приведенное уже Муавром (1718/1738, с. 6), см. в начале § 9.

8. Таблица семизначных логарифмов François Callet. Paris, 1795.

9. Аналогичное утверждение высказал Лаплас (1776/1891, с. 152; 1814/1999, столбец 837 левый).

10. Муавр (1718/1756, с. 37 – 38) уже указал это. Здесь и во многих других местах, следуя давнишней традиции, Пуассон приводит числовые данные с явно избыточными значащими цифрами.

11. См. Прим. № 7.

12. См. конец § 8.

13. Пуассон рассматривал азартные игры, но ввёл искусность игроков!

14. Более точно, сложным событием следовало назвать E + (F + G + …), так что F′ = F без E.

15. Заметим, что $a_0 + a_1t + \ldots + a_mt^m$ называется производящей функцией последовательности $\{a\}$.

16. Следует достаточное условие.

17. После μ тиражей, в которых появилось m белых шариков и n чёрных, шанс наступления белого шарика в новом тираже будет зависеть от этих чисел и окажется равным a'/c'. Однако, для того, кто только знает, что из урны извлекли μ шариков, вероятность появления белого шарика в новом тираже будет весьма отличной от указанного шанса. В соответствии с посланной мне заметкой Мондезира, бывшего студента Политехнической школы, эта вероятность не зависит от m и n и, как и до тиражей, равна a/c.

Для подтверждения этого положения предположим, что

$$a = 4,\ b = 3,\ c = 7,\ \mu = 2,\ c' = 5.$$

По отношению к m и n имеются три возможных не равновероятных случая, а именно $m = 2$, $n = 0$; $m = 1$, $n = 1$; $m = 0$, $n = 2$. Их вероятности определяются из выражения для П и равны соответственно 2/7, 4/7 и 1/7, а шансы появления белого шарика в новом тираже оказываются равными 2/5, 3/5 и 4/5. По правилам §§ 5 и 10 полная вероятность указанного события равна сумме произведений

$$\frac{2}{7}\cdot\frac{2}{5}+\frac{4}{7}\cdot\frac{3}{5}+\frac{1}{7}\cdot\frac{4}{5}=\frac{4}{7}=\frac{a}{c}.$$

Общее доказательство см. в заметке Мондезира, который предполагает опубликовать её в […]. Сформулированное положение очевидно, если $a = b$, потому что в этом случае, либо до, либо после получения этого результата, у того, кто о них ничего не знает, нет никаких оснований полагать, что в новом тираже появится белый шарик, а не чёрный. Вероятность каждого из этих исходов неизменно равна 1/2. Можно также заметить, что, при бесконечных a и b, это положение соответствует другому, которое будет доказано ниже. Именно, оказывается достоверным, что числа m и n остаются в том же соотношении, как и a к b, и можно быть уверенным, что $a':b' = a:b$, так что шанс и вероятность появления нового белого шарика не отличаются друг от друга и равны $a':c' = a:c$. Автор

Как и в начале § 11, Пуассон фактически рассматривал субъективные вероятности. См. Шейнин (2002). Статью Мондезира мы указали в Библиографии. О. Ш.

18. Нужно ли было доказывать это утверждение?

19. Вычисления Лапласа были исключительно сложными. Он и ранее, в 1774 и 1786 гг., решал эту задачу. Муавр (1712/1984, Задачи 18 и 19; 1718/1756, Задача 39) исследовал появление любых исходов любого числа игральных костей с произвольным числом граней.

20. Те же бесполезные системы игры упоминал уже Монмор (Montmort 1708, Préface).

21. Лучше всех высказался Бертран (Bertrand 1888, с. XXII): рулетка *не имеет ни совести, ни памяти*.

22. Общедоступными (publics), как можно понять, Пуассон назвал игры, регулярно проводимые в официально существовавших игорных домах. См. также Прим. 12 к Предисловию.

23. Этот параграф был написан до того, как последний вышедший закон о финансах запретил азартные игры с 1 января 1838 г. Автор

24. Термин *ожидание* Лаплас (1812/1886, с. 189) заменил на *математическое ожидание*, чтобы отличить его от ставшего модным *морального ожидания* (см. ниже). От этого уточнения, укоренившегося во французской и русской литературе, давно уже следовало отказаться.

25. Пояснение Пуассона явно недостаточно.

26. Следует доказательство по индукции (притом нестрогое), но нужно ли оно?

27. Описанную *петербургскую игру* придумал Николай Бернулли (Монмор 1708/1713, с. 402), а общеизвестной она стала после того, как Даниил Бернулли (1738) предложил разрешить её парадокс при помощи морального ожидания. Свой мемуар он опубликовал в Петербурге. Пуассон рассмотрел эту игру недостаточно. Кондорсе и Лакруа заметили, что даже её возможно бесконечный вариант соответствует одному-единственному испытанию и что для её исследования необходимо рассматривать множество таких игр. Фрейденталь (Freudenthal 1951) предложил кроме того, что роли игроков должны в каждой игре определяться по жребию.

28. Эту проблему рассматривал Лаплас в своих ранних мемуарах и позже (1812, гл. 7).

Библиография

Гнеденко Б. В. (1950), *Курс теории вероятностей.* М., 1954.

Курно О. (1843, франц.), *Основы теории шансов и вероятностей*. М. Перевод Н. С. Четверикова

Чебышев П. Л. (1845), Опыт элементарного анализа теории вероятностей. *Полн. собр. соч.*, т. 5. М. – Л., 1951, с. 26 – 87.

Шейнин О. Б., Sheynin O. (2002), Sampling without replacement: history and applications. *Z. f. Geschichte u. Ethik d. Naturwissenschaften, Technik u. Medizin*, Bd. 10, pp. 181 – 187.

Бернулли Д. (1738, латин.), Опыт новой теории измерения жребия. В сборнике *Теория потребительского поведения и спроса*. СПБ, 1999, с. 11 – 27.

Bertrand J. (1888), *Calcul des probabilités*. Paris, 1907, New York, 1970, 1972.

De Moivre A. (1712, латин.), De mensura sortis, or The measurement of chance. *Intern. Stat. Rev.*, vol. 52, 1984, pp. 236 – 262. Commentary by A. Hald on pp. 229 – 236.

--- (1718), *Doctrine of Chances*. London, 1738, 1756. Reprint of last edition: New York, 1967.

Freudenthal H. (1951), Das Petersburger Probleme im Hinblick auf Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Math. Nachr.*, Bd. 4, pp. 184 – 192.

Laplace P. S. (1776), Recherches sur l'intégration des équations différentielles aux différences finies. *Oeuvr. Compl.*, t. 8. Paris, 1891, pp. 69 – 197.

--- (1812), *Théorie analytique des probabilités*. Ibidem, t. 7, No. 1. Paris, 1886.

--- (1814, фпанц.), *Опыт философии теории вероятностей*. В книге Прохоров Ю. В., редактор (1999), *Вероятность и математическая статистика. Энциклопедия*. М., с. 834 – 863.

Mondesir E. (1837), Solution d'une question qui se présente dans le calcul des probabilités. *J. math. pures et appl.*, t. 2, pp. 3 – 10.

Montmort P. R. (1708), *Essay d'analyse sur les jeux de hazard*. Paris, 1713, New York, 1980.
Poisson S.-D. (1825 – 1826), Sur l'avantage du banquier au jeu de trente-et-quarante. *Ann. math. pures et appl*., t. 16, pp. 173 – 208.

Глава 2. Общие правила, продолжение. Вероятности причин и будущих событий, выведенные из наблюдений прошедших событий

Опечатки/ошибки, не замеченные автором

1. В § 34, с. 95 оригинала, строка 2 снизу. Вероятность p_1 следует заменить на вероятность p_n.

2. В § 37, с. 101 оригинала, в приведенной единой строке формул для y_1 и y_2, в формуле для y_2, во втором слагаемом знаменателя ошибочно указано y_2 вместо y_1. Там же, при переходе от y_x к y_{x-1} x был заменён на $(x - 1)$, а не на $(y - 1)$, как в тексте.

3. В § 38, с 104 оригинала. В строке 10 второго абзаца вместо n напечатано n'.

Там же, с. 106 оригинала, в приведенной единой строке для w_n и w_i в формуле для w_i, в первом слагаемом знаменателя, ошибочно указано $q_n p_n$ вместо $q_i p_i$.

4. В § 39, с. 109 оригинала, строка 6. Число шариков должно быть a_n и a_i, а не a_n и a_x.

5. В § 40, с. 113 оригинала. В первой строке после выключенной формулы ссылка должна быть на § 38, а не на § 39.

6. В § 45, с. 124 оригинала, в конце параграфа дана ссылка на § 27 вместо § 26.

7. В § 46, с. 125 оригинала в подынтегральной функции $x^m(1 - x)^n$ ошибочно указан множитель $(1 - x)^m$.

8. В § 53, с. 141 оригинала. Доказывая *третье основное положение*, автор указал, что величина А принимает μ, а не λ значений.

9. В § 55, с. 145 оригинала, при описании извлечения шариков из урн B_1, B_2, …, B_μ указано, что число тиражей равно m а не μ тиражей.

10. В § 56, с. 149 оригинала. Два неравенства повторены вместо того, чтобы при повторении изменить их знаки.

11. В § 60, с. 156 оригинала. Сразу же над последней выключенной формулой $(\alpha + u)$ должно быть вставлено вместо z, а не вместо 2.

Все исправления внесены в перевод.

27. В предыдущей главе рассмотренные нами правила предполагали известными шансы некоторых событий и имели целью вывести вероятности других событий, составленных из первых[1]. Здесь мы описываем правила для вычисления вероятностей причин по наблюдённым событиям, а затем вероятностей будущих событий. Но вначале удобно пояснить точный смысл слова *причина*, который не совпадает с тем, что понимается в обычном словоупотреблении.

Когда говорят, что некоторая вещь является *причиной* другой, первой обычно приписывают *возможность* наверняка произвести вторую, но не желают тем самым выразить ни суть её могущества, ни метода её действия. В конце этой главы мы вернёмся к этому понятию *причинности*, но пока нам достаточно сказать, что в исчислении вероятностей слово *причина* понимается более широко. *Причина* С относительно некоторого события Е рассматривается как что-то, придающее появлению Е определённый и свойственный ему шанс[2]. В обычном словоупотреблении С – причина этого шанса, а не самого события, и действительное наступление Е имеет место при сочетании С с другими причинами или обстоятельствами, которые не влияют на свойственный этому событию шанс.

Пусть этот известный или неизвестный шанс равен p; он, вообще, отличен от соответствующей вероятности. В то же время причина С придаёт шанс $1 - p$ противоположному событию F. Если $p = 1$, то причина С наверняка приведёт к Е, и таким образом окажется его причиной в собственном смысле. Если же $p = 0$, она окажется такой же причиной для F. Множество причин, которые сочетаются для наступления события, но не оказывают влияния на величину его шанса, т. е. на отношение числа случаев, благоприятных его появлению, ко всем возможным случаям, это то, что мы должны понимать под *случайностью* (hasard)[3].

Например, при игре в кости событие, происходящее каждый раз, является следствием числа граней и возможной нерегулярности формы и плотности костей, а также встряхиваний перед бросками. Эти встряхивания являются причинами, которые никак не влияют на появление некоторой грани, а нужны они для того, чтобы исключить влияние положения костей в их стаканчике до броска, чтобы их исходное положение не было известно игрокам. Если эта цель достигнута, относительные шансы появления граней будут зависеть лишь от их числа и недостатков костей, которые могут привести к неравенству этих шансов. Мы говорим, что нечто произошло случайно, если при этом не изменились относительные шансы наступления возможных событий. Пусть из урны, содержащей белые и чёрные шарики, случайно, не учитывая их расположения и полагая, что у всех один и тот же диаметр[4], извлекли шарик. Тогда шанс появления белого шарика будет, очевидно, зависеть лишь от количества тех и других, и доказывается, что он равен отношению числа белых шариков к их общему числу.

Причина С может быть физической или моральной. При игре в орлянку физическое строение подбрасываемой монеты приводит, вообще говоря, к немного различным шансам появления возможных исходов, а при судебном рассмотрении уголовных дел шансы верного и ошибочного голосования каждого присяжного определяются его моральными качествами, т. е. его способностями и честностью в данном деле.

Иногда причина С возникает при сочетании морального и физического. Так, в каждом виде измерений или наблюдений шанс погрешности каждой величины зависит от умения наблюдателя и более или менее совершенного устройства его инструмента[5]. Но во всех случаях в исчислении вероятностей различные причины учитываются вне зависимости от их сути, а только по отношению к величине производимых ими шансов. Это исчисление поэтому равным образом применимо к моральным и физическим вещам. В большинстве проблем шанс, происходящий от заданной причины С, заранее неизвестен, порой неизвестна и сама причина события или её шанс.

Если этот последний постоянен, его определяют, как будет сказано ниже, по достаточно длинной серии испытаний. Но, например, при голосовании присяжных шанс ошибочного решения изменяется от одного из них к другому, и несомненно у одного и того же присяжного в зависимости от вида уголовного дела. Повторение испытаний исключено, но можно выводить из наблюдений, как будет сказано ниже, не шанс ошибки, присущей данному присяжному, а лишь некоторую вероятность всех присяжных данного ассиза. Этого, однако, достаточно для решения задач, которые являются специальной темой моего сочинения.

Часто существует много различных причин, которые могут в сочетании со случаем привести к появлению данного события Е или противоположного события F. Перед их наступлением каждая такая причина имеет некоторую вероятность, которая изменяется после наблюдения Е или F. Если предположить, что известно, какой шанс каждая из этих возможных причин, действуй она достоверно, придаёт появлению Е или F, мы вначале определим вероятности всех этих причин после наблюдения, а затем и вероятность любого другого будущего события, зависящего от них.

28. Пусть наблюдалось именно событие E, наступление которого может быть приписано некоторому числу m различных, равновероятных до наблюдения и взаимоисключающих причин и только им. Появление E изменяет вероятности этих возможных причин, и их-то требуется теперь определить. Для этого служит следующая теорема.

Вероятность каждой возможной причины наблюдаемого события равна вероятности, которую она придаёт событию, будь она достоверна, делённую на сумму вероятностей этого события, вытекающих от всех возможных причин, приводящих к нему. Итак, пусть

$$C_1, C_2, \ldots, C_m$$

будут m возможными причинами наступления события E и

$$p_1, p_2, \ldots, p_m$$

известными вероятностями его появления относительно указанных различных причин: p_n – вероятность наступления E, будь C_n единственной причиной. Иначе говоря, будь C_n достоверно, это исключило бы все остальные причины. Далее, пусть

$$w_1, w_2, \ldots, w_m$$

будут неизвестными вероятностями этих же причин, так что w_n – вероятность причины C_n; другими словами, вероятность того, что наступление E было вызвано причиной C_n. Требуется доказать, что

$$w_n = \frac{p_n}{p_1 + p_2 + \ldots + p_n + \ldots + p_m}.$$

Каково бы ни было событие E, его можно уподобить появлению белого шарика из урны, содержащей шарики белого и чёрного цветов. Пусть имеется m таких урн с соотношением p_n белых шариков ко всем шарикам в урне A_n. Урна выбирается случайно и каждая представляет собой одну из причин появления белого шарика, т. е. урна A_n соответствует причине C_n. Требуется определить вероятность того, что шарик был извлечён из A_n.

Приведём все дроби p_n к общему знаменателю:

$$p_1 = \frac{\alpha_1}{\mu}, p_2 = \frac{\alpha_2}{\mu}, ..., p_m = \frac{\alpha_m}{\mu},$$

причём и μ, и числители этих дробей целые числа. Шанс извлечения белого шарика из некоторой урны A_n не изменится, если обозначить их число в ней через α_n, а число всех шариков через μ; общее число шариков во всех урнах теперь таким образом одно и то же.

Объединим все шарики в единой урне А и придадим номер 1 тем, которые были в урне A_1, номер 2 – бывшим в урне A_2 и т. д. В соответствии с леммой § 10 вероятность w_n того, что белый шарик, извлечённый из А, раньше был в урне A_n, окажется той же, как если бы он появился из урны А с номером *n*. Эта вероятность равна отношению α_n к сумме $\alpha_1 + \alpha_2 + \ldots + \alpha_m$:

$$w_n = \frac{\alpha_n}{\alpha_1 + \alpha_2 + ... + \alpha_m}.$$

В силу предшествующих уравнений эта величина совпадает с выражением w_n, которое требовалось вывести.

29. При вычислении вероятностей многих последовательных событий необходимо учитывать возможное влияние появления одного из них на шанс последующего, см. § 9. Иногда кроме того при оценке этого шанса следует иметь в виду вероятности различных причин предыдущего события или различных способов, при помощи которых оно может произойти. Это будет видно, например, из следующей задачи.

Пусть даны *m* урн А, В, С, … с белыми и чёрными шариками, а шансы извлечения белого шарика из них равны соответственно *a, b, c*, … Из случайно выбранной урны извлечён шарик, затем второй из какой-то другой урны и т. д. Таким образом, после каждого тиража исключается урна, из которой был извлечён шарик. Требуется определить вероятность извлечения *n* белых шариков в *n* тиражах ($n \le m$). Примем для сокращения письма

$a + b + c + d + \ldots = s_1$,
$ab + ac + ad + bc + bd + cd + \ldots = s_2$,

$abc + abd + bcd + \ldots = s_3$, $abcd + \ldots = s_4$, …

Величина s_2 состоит из C_m^2 членов, s_3 – из C_m^3 членов и т. д. Вероятность извлечь белый шарик при первом тираже равна s_1/m. Если шарик, будь он белым или чёрным, был в урне А, то вероятность извлечь такой же шарик во втором тираже равна $(s_1 - a)/(m - 1)$ или $(s_1 - b)/(m - 1)$, если первый шарик был извлечён из урны В и т. д. Поэтому, в соответствии с правилами §§ 9 и 10,

$$\frac{1}{m-1}[\alpha(s_1 - a) + \beta(s_1 - b) + \gamma(s_1 - c) + \ldots]$$

окажется полной вероятностью появления белого шарика во втором тираже. Здесь α, β, γ, … – вероятности того, что для первого тиража были выбраны соответственно урны А, В, С, … Они не равны друг другу[6], и, в соответствии с § 28,

$$\alpha = a/s_1,\ \beta = b/s_1,\ \gamma = c/s_1,\ \ldots$$
$$a(s_1 - a) + b(s_1 - b) + c(s_1 - c) + \ldots = 2s_2.$$

Во втором тираже вероятность извлечения белого шарика окажется поэтому равной $2s_2/[(m - 1)s_1]$. Аналогично для третьего тиража вероятность равна $(s_1 - a - b)/(m - 2)$, если только для первых двух были выбраны урны А и В при любом цвете появившихся шариков. Если же выбранными оказались урны А и С, то эта вероятность будет равна $(s_1 - a - c)/(m - 2)$ и т. д. Поэтому полная вероятность исхода белого шарика при третьем тираже равна

$$\frac{1}{m-2}[g(s_1 - a - b) + h(s_1 - a - c) + k(s_1 - b - c) + \ldots].$$

Здесь g, h, k, … – вероятности, что для первых двух тиражей были выбраны урны А и В, А и С, В и С и т. д. В соответствии с § 28

$$g = ab/s_2,\ h = ac/s_2,\ k = bc/s_2,\ \ldots$$

и так же само

$$ab(s_1 - a - b) + ac(s_1 - a - c) + bc(s_1 - b - c) + \ldots = 3s_3$$

и потому вероятность выхода белого шарика в третьем тираже становится равной $3s_3/[(m-3)s_2]$. Это рассуждение можно без труда неопределённо продолжить. В результате оказывается, что вероятности извлечения белого шарика в каждом из n первых тиражах равны

$$\frac{s_1}{m},\ \frac{2s_2}{(m-1)s_1},\ \frac{3s_3}{(m-2)s_2},\ldots,\ \frac{ns_n}{(m-n+1)s_{n-1}},$$

а требуемая вероятность будет равна произведению этих n дробей (см. § 5), т. е. $(1/C_m^n)s_n$.

Значение этой вероятности можно подтвердить, заметив, что каждое произведение есть вероятность появления n белых шариков из n определённых урн, выбранных из A, B, C, … и что поэтому сумма всех этих произведений, разделённая на их число, есть вероятность извлечь n белых шариков из n этих случайно выбранных урн. Эту вероятность очевидно и следовало вычислить. При $n = m$ будет $C_m^m = 1$ и она равна s_m, что немедленно следует из правила § 5.

30. Далее, пусть E′ – другое, отличное от E событие, зависящее, однако, от тех же самых причин C_n. Обозначим через

$$p'_1, p'_2, \ldots, p'_m$$

шансы появления E′ относительно указанных различных причин: p'_n есть данная вероятность, что E′ наступит, если причина C_n достоверна. Эта причина – единственно вероятная, и её вероятность равна w_n.

Появление E′ есть составное событие, шанс которого равен произведению этих двух вероятностей (§ 5). Более того, полная вероятность E′ есть сумма шансов относительно m различных способов возможного появления этого события (§ 10), т. е. сумма произведений $p'_n w_n$, относящихся к m возможным причинам C_n событий E и E′:

$$w' = p'_1 w_1 + p'_2 w_2 + \ldots + p'_n w_n + \ldots + p'_m w_m.$$

Подставляя значения w_1, w_2, … мы получим

$$w' = \frac{p_1 p'_1 + p_2 p'_2 + \ldots + p_n p'_n + \ldots + p_m p'_m}{p_1 + p_2 + \ldots + p_n + \ldots + p_m}.$$

Такова формула для вычисления вероятностей будущих событий по наблюдению прошедших. Можно придти к этому же выражению без посредства общих причин E и E′, полагая последние составными событиями, зависящими от одного и того же простого события. Наше рассуждение равным образом применимо к этому другому методу рассмотрения вопроса, но при желании можно непосредственно вернуться к предыдущему.

Если E и E′ составлены из одного и того же события G, которое до появления E допускало различные равновероятные шансы

$$g_1, g_2, \ldots, g_n, \ldots, g_m,$$

их можно рассматривать как столько же различных причин появления E и E′. Полагая поэтому g_n причиной, которую мы раньше обозначали C_n, и вероятность g_n, как мы обнаружили, примет значение w_n, т. е. w_n будет вероятностью, что шанс G равен g_n. Предыдущее выражение для w' будет вероятностью появления E′ в результате m возможных значений шанса G. В этой формуле p_n и p'_n выражают данные вероятности появления E и E′, если g_n достоверно является шансом G.

31. Это определение вероятности E′ после появления E нельзя смешивать с каким-то влиянием наступления прошедших событий на появление будущих, что нелепо. Если я, к примеру, уверен, что урна A содержит 3 белых шарика и 1 чёрный, то для меня достоверно, что шанс извлечения белого равен 3/4. Следовательно, если E′ – выход двух белых шариков с возвращением первого, шанс этого события окажется равным $(3/4)^2 = 9/16$ вне зависимости от возможно наблюдённого E.

Если E – последовательный выход с возвращением некоторого числа белых шариков и какого-то числа чёрных, я всегда должен буду вне зависимости от соотношения этих чисел ставить 9 против 7 за появление E′. Но если шанс простого события G мне неизвестен, и я только знаю, что он может принимать лишь определённые значения, наблюдение E позволит мне определить вероятность каждого из них, а потому и вероятность E′. Это наблюдение усилит или ослабит моё бывшее основание верить, что появится E′, но нисколько не повлияет на это будущее событие или на свойственный ему шанс. Для кого-то, кто наблюдает другое событие E_1, зависящее от того же простого события G, основание верить в появление E′ может быть намного сильнее или слабее, чем у меня, однако это никак не изменит шанс, свойственный E′.

Относительно двух лиц, один из которых наблюдал событие E, а другой – E_1, составленные из одного и того же события G, не следует забывать, что если E_1 включает E и что-то ещё, мнение второго о появлении нового события E′, также зависящего от G, окажется более обоснованным и его следует предпочесть (§ 1). Если предположить, что наблюдение E_1 приводит к вероятности k будущего события E′, а наблюдение E – к вероятности h того же события, то второе лицо более обоснованно может ставить k против $(1 - k)$, чем первый – h против $(1 - h)$ за появление E′, каковы бы ни были дроби h и k, большие или меньшие половины и каков знак разности $(h - k)$.

32. Прежде, чем идти дальше, уместно привести несколько простых примеров на применение предыдущих выражений для w_n и w', которые мы вкратце запишем в виде

$$w_n = \frac{p_n}{\sum p_n},\ w' = \frac{\sum p_n p'_n}{\sum p_n}.$$

Здесь $\sum$ означает сумму m значений индекса n от $n = 1$ до m.

Известно, что урна B содержит m белых и чёрных шариков. После появления белого шарика требуется определить вероятность, что в ней находилось n таких шариков. По поводу их числа можно сформулировать m различных взаимоисключающих гипотез: урна содержала m белых шариков; $(m - 1)$ белых и 1 чёрный; …; 1 белый и $(m - 1)$ чёрных. Все они равновозможны, и их можно принять за причины C_1, C_2, … события E, т. е. извлечения белого шарика.

Предположим, что из m шариков n белых, тогда вероятность указанного события будет равна $p = n/m$ и

$$\sum p_n = \frac{1}{2}(m+1),\ \text{и}\ w_n = \frac{2n}{m(m+1)}$$

есть вероятность, что урна действительно содержала n белых шариков. Она может быть равна 1/2 только при $m = n = 3$. Вообще, после выхода белого шарика вероятность того, что в урне находятся только белые шарики, или что $n = m$, равна

$2/(m + 1)$. Если E′ – извлечение ещё одного белого шарика, то вероятность w' будет зависеть от того, был ли он возвращён или нет. В первом случае

$$p'_n = p_n = \frac{n}{m}, \ \sum p_n p'_n = \frac{1}{m^2} \sum n^2.$$

Однако

$$\sum \frac{n(n+1)}{1 \cdot 2} = \frac{m(m+1)(m+2)}{3!}, \ \sum n = \frac{m(m+1)}{1 \cdot 2}$$

и потому

$$\sum n^2 = 2\sum \frac{n(n+1)}{1 \cdot 2} - \sum n = \frac{m(m+1)(2m+1)}{3!}, \ w' = \frac{2m+1}{3m}.$$

Во втором случае число белых и всех шариков в урне уменьшилось после первого тиража на единицу, и потому

$$p'_n = \frac{n-1}{m-1}, \ \sum p_n p'_n = \frac{1}{m(m-1)} \sum n(n-1).$$

Но неизменно $p_n = n/m$, $\sum p_n = (m + 1)/2$ и $w' = 2/3$, поскольку

$$\sum \frac{n(n-1)}{1 \cdot 2} = \frac{(m-1)m(m+1)}{3!}.$$

Вероятность извлечь из урны белый шарик после того, как из неё вышел без возврата такой же, поэтому не зависит от числа m белых и чёрных шариков в урне и всегда равна 2/3. Значение w' в первом случае также уменьшается до 2/3, как и должно было быть, если m – очень большое число, рассматриваемое как бесконечность. Если известно, что из исходного числа m извлечено $(m - 1)$ белых шариков, вероятность последнему быть белым равна $m/(m + 1)$. Можно сформулировать только две гипотезы, C_1 и C_2, а именно, что все m шариков были белые и что один шарик был чёрный. В первом случае наблюдённое событие достоверно; во втором его вероятность, т. е. шанс извлечь $(m - 1)$ белых шариков, совпадает с вероятностью того, что в урне остался чёрный шарик. И, как и вначале, этот шарик может быть любым из m бывших в урне, и потому указанная вероятность равна $1/m$. Тогда

$$p_1 = 1,\ p_2 = \frac{1}{m},\ \text{и}\ w_1 = \frac{p_1}{p_1 + p_2} = \frac{m}{m+1}$$

будет вероятностью первой гипотезы. Случай $m = 1$ не включен в это значение w_1, которое оказывается равным 1/2, что было ясно с самого начала. [Здесь принята во внимание поправка автора.]

33. Вот ещё одно более непосредственное приложение предыдущих формул, при котором полное число белых и чёрных шариков в урне неизвестно, известно лишь, к примеру, что оно не более трёх. Зарегистрированным событием E является появление x белых шариков в n тиражах с возвращением. Если $x \neq 0$ и $\neq n$, можно сформулировать только три гипотезы: 1 белый шарик и 1 чёрный (C_1); 2 белых и 1 чёрный (C_2); 2 чёрных и 1 белый (C_3). Соответствующие вероятности события E равны

$$p_1 = (1/2)^x(1/2)^{n-x} = 1/2^n,\ p_2 = (2/3)^x(1/3)^{n-x} = 2^x/3^n,$$
$$p_3 = (1/3)^x(2/3)^{n-x} = 2^{n-x}/3^n.$$

Положим для краткости

$$3^n + 2^{n+x} + 2^{2n-x} = \mu,$$

тогда $w_1 = 3^n/\mu$, $w_2 = 2^{n+x}/\mu$, $w_3 = 2^{2n-x}/\mu$ будут соответствовать вероятностям гипотез. За E′ примем будущее событие, извлечение нового белого шарика. Его вероятности окажутся равными $p'_1 = 1/2$, $p'_2 = 2/3$, $p'_3 = 1/3$, а полная вероятность

$$w' = \frac{(1/2)3^n + (2/3)2^{n+x} + (1/3)2^{2n-x}}{3^n + 2^{n+x} + 2^{2n-x}}.$$

В случае $n = 2x$

$$w_1 = \frac{9^x}{9^x + 2\cdot 8^x},\ w_2 = w_3 = \frac{8^x}{9^x + 2\cdot 8^x},\ w' = \frac{(1/2)9^x + 8^x}{9^x + 2\cdot 8^x} = \frac{1}{2}.$$

Значение w' и должно было быть равно 1/2, потому что было извлечено поровну белых и чёрных шариков и нет основания верить, что белый шарик появится в новом тираже скорее, чем чёрный. Тем не менее, 9^x должно было бы быть более чем вдвое больше, чем 8^x, т. е. x должен был бы быть более пяти, чтобы можно было ставить боле одного к одному за то, что числа белых и чёрных шаров совпадают или что урна содержит их по одному. Если x – очень большое число, вероятность этой гипотезы очень мало отличается от достоверности.

Если i – целое число, $x = 2i$ и $n = 3i$, то

$$w' = \frac{(1/2)27^i + (2/3)32^i + (1/3)16^i}{27^i + 32^i + 16^i}.$$

При очень больших i эта величина очень мало отличается от 2/3. В то же время вероятность w_2 того, что урна содержит 2 белых шарика и 1 чёрный, также очень немного отличается от достоверности.

Допустим ещё, что $n = 3x$. Тогда

$$w' = \frac{(1/2)27^x + (2/3)16^x + (1/3)32^x}{27^x + 16^x + 32^x}.$$

При очень больших x эта вероятность окажется весьма близкой к 1/3. Вероятность w_3 того, что урна В содержит 1 белый и 2 чёрных шариков также очень близка к достоверности.

В третьем случае, в предположении, что число тиражей очень велико, значение вероятности w' появления нового белого шарика существенно приближается к отношению числа извлечённых белых шариков к общему числу тиражей, и неизменно это отношение с вероятностью, очень близкой к достоверности, равно отношению числа белых шариков в урне ко всему их числу, т. е. к шансу, свойственному извлечению белого шарика.

Ниже будет указано, что если событие какой-либо природы наблюдается некоторое число раз в очень длинной серии испытаний, отношение первого числа ко второму будет весьма вероятным значением этого события, притом очень близким к его известному или неизвестному шансу. В нашем примере окажется, что этот шанс может быть только равен 1/2, 2/3 или 1/3, т. е. x/n = 1/2, 2/3 или 1/3 – также единственные, которые можно будет предположить правдоподобными, если только x и n – очень большие числа.

34. В предыдущем изложении мы предположили, что до появления Е все причины C_1, C_2, …, которые могут быть ему приписаны, были равновозможными. Но если заранее известно, что имеется какое-то основание верить, что одна из них существует скорее, чем другая, то до наблюдения это неравенство шансов должно быть учтено при оценке тех вероятностей указанных причин, которые они приобретут после появления Е. Эта необходимость является существенным обстоятельством в теории вероятностей[7], особенно, как мы разъяснили в Предисловии, в вопросах о решениях трибуналов.

Впрочем, доказательство в § 28 легко обобщить на общий случай, в котором причины события Е до наблюдения обладали некоторыми известными вероятностями. Как и в том параграфе, заменим событие Е извлечением белого шарика из какой-либо урны A_1, A_2, … и вначале предположим, что существует равная возможность выбора любой их них. Вероятность того, что он вышел из некоторой урны A_n равна $p_n/\sum p_n$, где p_n, как и всегда, является отношением числа белых шариков в урне A_n ко всему числу шариков в ней, а сумма $\sum$ распространяется на все урны. По правилу § 10 [полная] вероятность извлечения белого шарика равна сумме подобных дробей.

Представим себе, что урны A_1, A_2, … составлены из некоторого числа a_1 урн A_1 с равными отношениями p_1, a_2 урн A_2, с равными p_2, … и, наконец, a_i урн A_i с равными p_i. Число всех урн

$$s = a_1 + a_2 + \ldots + a_i,$$

где i – число групп урн одного и того же состава. Суммы [см. выше] можно заменить на $\sum a_n p_n$ по всем группам или по всем индексам n от 1 до i. Вероятность извлечения белого шарика из группы n будет равна

$$w_n = \frac{a_n p_n}{\sum a_n p_n}.$$

До наблюдения вероятность извлечения белого или чёрного шарика из этой группы равнялась $a_n/s = q_n$, так что $a_n = sq_n$ и

$$w_n = \frac{q_n p_n}{\sum q_n p_n}.$$

Опять же, до наблюдения различные группы представляли собой все возможные не равновероятные причины C_1, C_2, … числом i, события E. Дробь q_n выражает предварительную вероятность того, что E будет вызвано причиной C_n, а w_n есть вероятность, что оно было вызвано именно этой причиной. Поскольку причины C_1, C_2, … взаимоисключающи, q_n и w_n являются вероятностями существования этой причины до и после наблюдения. Выражение для w_n таким образом указывает, что вероятность каждой из этих возможных причин наблюдённого события равна произведению вероятности q_n этой причины до наблюдения и вероятности p_n, которую она приписывает этому событию, если оно достоверно, разделённому на сумму […].

Вероятность w' будущего события E′, зависящего от тех же причин, равна, подобно сказанному выше,

$$w' = \sum w_n p'_n = \frac{\sum q_n p_n p'_n}{\sum q_n p_n}.$$

Пусть E′ произошло после E, а E″ – третье событие, зависящее от тех же причин, и p''_n – шанс, доставляемый причиной C_n, если она достоверна, будущему появлению E″. Вероятность этой причины после наблюдения E и до наступления E′ равна w_n, а после этого, по указанному выше правилу, она становится равной

$$\frac{w_n p'_n}{\sum w_n p'_n} = \frac{q_n p_n p'_n}{\sum q_n p_n p'_n}.$$

Умножая её на p''_n, мы получим вероятность появления E″, вызванного причиной C_n, а полная вероятность этого события равна

$$w'' = \frac{\sum q_n p_n p'_n p''_n}{\sum q_n p_n p'_n}.$$

Это выражение, как и в случае w', можно вывести, заменяя p'_n на p''_n и p_n на $p_n p'_n$, а по отношению к причине C_n произведение $p_n p'_n$ есть шанс наблюдённого события, т. е. последовательности событий E и E′.

35. Вот очень простой пример на предыдущее правило, который может служить для проверки его точности и необходимости. Пусть на столе находятся две карты неизвестного цвета. Перевернув одну из них, мы видим, что она красная. Об их цвете можно сформулировать только две гипотезы: обе – красные; и по одной красной и чёрной. Если происхождение карт совершенно неизвестно, то до наблюдения эти две возможные причины наблюдённого события были равновероятны. После наблюдения, как и одном из примеров § 32, вероятность первой гипотезы становится равной 2/3. Можно ставить 2 против одного за то, что и вторая карта – красная.

Но пусть известно, что, к примеру, эти карты были выбраны наугад из колоды для игры в *пикет* из 16 красных и 16 чёрных карт. Тогда до наблюдения (§ 18) вероятности этих гипотез окажутся равными $q_1 = 16{\cdot}15/32{\cdot}31$ и $q_2 = 2{\cdot}16{\cdot}16/32{\cdot}31$ и $p_1 = 1$, $p_2 = 1/2$, а потому после наблюдения

$$w_1 = \frac{q_1 p_1}{q_1 p_1 + q_2 p_2} = \frac{15}{16}$$

окажется вероятностью первой гипотезы. Таким образом, вместо 2:1 соотношение шансов гипотез теперь равно всего лишь 15:16.

Это значение w_1 может быть сразу же проверено, так как ясно, что можно задать поставленный вопрос иначе: выньте красную карту из всей колоды, и определите вероятность вынуть такую же карту из оставшихся, среди которых только 15 красных из 31. Вообще, пусть имеется m карт, a красных и b чёрных. Из них наугад выбрали n, причём среди $(n - 1)$ из отобранных оказалось a' красных и b' чёрных. По предыдущему правилу, вероятности последней из отобранных быть красной или чёрной равны соответственно

$$w_1 = \frac{a - a'}{m - n + 1},\ w_2 = \frac{b - b'}{m - n + 1}.$$

Как оно и должно было быть, w_1 является также вероятностью выбрать красную карту из первоначальной колоды, уменьшенной до $(m - n + 1)$ карт с $(a - a')$ красными после извлечения из неё a' красных и b' чёрных карт. Так же можно проверить w_2, ибо $w_1 + w_2 = 1$, поскольку $a + b = m$ и $a' + b' = n - 1$.

36. Общее следствие из правила § 34 состоит в том, что если событие E и будущее событие E′ зависят от одной и той же причины, вероятность E′ зависит не только от наблюдения E. При её определении следует учесть возможное знание об их общей причине, существовавшее до наблюдения,. Вероятность события

E′ может оказаться различной для двух лиц, которые наблюдали появление E, но до этого имели различные сведения о поставленном вопросе.

Аналогично обстоит дело при сомнениях или критичности, также относящихся к исчислению вероятностей (§ 3). Когда требуется установить, произошло или нет засвидетельствованное происшествие, необходимо учитывать не только шанс ошибки свидетеля, но и наше предшествовавшее знание. Пусть p – вероятность, что свидетель не обманывает нас ни невольно, ни умышленно и q – вероятность истинности происшествия до свидетельства о нём. Тогда после этого свидетельства его вероятность будет зависеть от p и q и определяется следующим образом.

Наблюдённое происшествие засвидетельствовано, но вовсе не является неоспоримым. Если оно произошло, то свидетель не обманывал нас, и p – вероятность происшествия. Она же оказывается равной $1 - p$ в предположении, что происшествие не наступило, потому что свидетель обманул нас. Перед свидетельством вероятностью первого предположения является q, и $1 - q$ – вероятностью второго. Пусть r будет вероятностью первого предположения, т. е. наступления происшествия после свидетельства о нём, тогда по правилу § 34

$$r = \frac{pq}{pq + (1-p)(1-q)},\ r - q = \frac{q(1-q)(2p-1)}{pq + (1-p)(1-q)},$$

что доказывает, что знаки $(r - q)$ и $(p - 1/2)$ совпадают. Это означает, что свидетель повышает или понижает прежнюю вероятность истинности происшествия в соответствии с нашим предположением о $p > 1/2$ или $p < 1/2$. Эта вероятность остаётся прежней, если, при $p = 1/2$, свидетель ничего не изменяет в предшествовавшей вероятности, и мы можем держать пари на равных о том, что свидетель говорит правду или нет. Если вначале у нас не было никаких оснований верить скорее в истинность или ложность свидетельствуемого происшествия, вероятность $q = 1/2$ и $r = p$. В этом случае вероятность истинности уже зависит только от правдивости и познаний свидетеля.

Нельзя предполагать, что из двух чисел, p и q, одно равно единице, а другое – нулю, но если p существенно приближается к достоверности, а q – ещё более к невозможности, так что $q/(1 - p)$ становится очень малой дробью, вероятность r будет тоже очень низка и почти равна ей. Так происходит, если происшествие противоречит общим законам природы, а свидетель его истинности сам по себе заслуживал бы серьёзного доверия.

Эти общие законы являются для нас результатом длительного опыта и потому обладают, если не абсолютной достоверностью, по меньшей мере очень высокой вероятностью, к тому же подкрепляемой той согласованностью, которую они представляют и которую не может уравновесить никакое противоположное свидетельство. Если всё же свидетельствуемое происшествие противоречит этим законам, то исходная вероятность, что оно очень точно описано, будет почти равна нулю. Пусть даже свидетель честен, достаточно, чтобы он не был совершенно непогрешим, чтобы шанс ошибки $1 - p$ оказался очень большим по сравнению с этой начальной вероятностью q, так что вероятность r после свидетельства можно было бы ещё считать неощутимой. В подобных случаях разумно было бы отрицать [даже] своё собственное свидетельство и полагать, что мы были обмануты нашими чувствами.

37. Предположим, что происшествие, вероятность которого мы рассматриваем, засвидетельствовано и вторым лицом. Пусть p' будет вероятностью, что оно не обманывает нас, и r' – вероятность истинности происшествия, вытекающая из двойного свидетельства. Заметив, что вне зависимости от нового обстоятельства эта истинность уже имела вероятность r, можно заключить, что выражение для r' должно выводиться из r при замене p и q на p' и r, так что

$$r' = \frac{p'r}{p'r + (1-p')(1-r)} = \frac{qpp'}{qpp' + (1-q)(1-p)(1-p')}.$$

Пусть второй свидетель противоречит первому, который утверждал, что происшествие имело место. Независимо от этого нового обстоятельства вероятность противоположного, т. е. ложного первого свидетельства, уже имела значение $1 - r$. Обозначим через r_1 вероятность ложности происшествия, вытекающую из обоих противоречащих друг другу свидетельств. Её можно вывести из r § 36, заменив p и q на p' и $(1 - r)$, так что

$$r_1 = \frac{p'(1-r)}{p'(1-r) + r(1-p')} = \frac{p'(1-p)(1-q)}{p'(1-p)(1-q) + qp(1-p')}.$$

Если $p = p'$, то r_1 сводится к $(1 - q)$. Противоречащие друг другу свидетельства одного и того же достоинства уничтожают друг друга, и вероятность ложности происшествия должна оставаться прежней. Можно без труда определить вероятность истинности или ложности происшествия, о котором высказаны противоположные суждения нескольких свидетелей. Но пусть все они единодушно утверждают, что происшествие имело место. Тогда вероятность, что оно истинно, приобретает следующую форму.

Пусть до этих свидетельств q было вероятностью истинности происшествия и y_x и y_{x-1} – её новые значения после x и $(x - 1)$ свидетельств, а $p^{(x-1)}$ – вероятность, что другой свидетель не обманул нас, утверждая, что происшествие истинно. Тогда y_x можно вывести из выражения для r § 36 при замене p и q на $p^{(x-1)}$ и y_{x-1}:

$$y_x = \frac{p^{(x-1)} y_{x-1}}{p^{(x-1)} y_{x-1} + (1 - p^{(x-1)})(1 - y_{x-1})},$$

причём y_0 есть первоначальное значение вероятности q. Полагая $x = 1, 2, 3, \ldots$, можно вывести следующие формулы

$$y_1 = \frac{pq}{pq + (1-p)(1-q)}\ , \ y_2 = \frac{p'q}{p'y_1 + (1-p')(q - y_1)}, \ldots$$

т. е. вычислить y_2, исходя из y_1, y_3 – исходя из y_2 и т. д. Примем для сокращения письма

$$\frac{1 - p^{(x-1)}}{p^{(x-1)}} = \rho_x,$$

так что предыдущее разностное уравнение 1-го порядка и его общее решение приобретут вид

$$y_x = \frac{y_{x-1}}{y_{x-1} + \rho_x (1 - y_{x-1})} = \frac{c}{c + (1-c)\rho_1 \rho_2 \ldots \rho_x}.$$

Здесь c – произвольная постоянная. Приняв $(x - 1)$ вместо x, окончательно получаем

$$y_{x-1} = \frac{c}{c + (1-c)\rho_1\rho_2...\rho_{x-1}},\ 1 - y_{x-1} = \frac{(1-c)\rho_1\rho_2...\rho_{x-1}}{c + (1-c)\rho_1\rho_2...\rho_{x-1}}.$$

Совместно с y_x эти значения переводят заданное уравнение в тождество. Константу c определяют по частному значению y_x, например, при $x = 0$. Соответственно, принимая значение произведения x сомножителей $\rho_1\rho_2 \ldots\rho_x$ равным единице, мы получаем $y_0 = q = c$. Для некоторого числа x свидетелей

$$y_x = \frac{q}{q + (1-q)\rho_1\rho_2...\rho_x}$$

и для свидетеля i величина ρ_i равна отношению вероятностей того, что он обманывает и не обманывает нас, так что $\rho_i > 1$ или < 1, если первая вероятность выше или ниже второй и $\rho = 1$, если они равны друг другу.

Если свидетелей очень много и, можно сказать, бесконечно много, и если $\rho_i > 1$ для всех из них, вероятность y_x истинности происшествия, которое они засвидетельствовали, равна нулю при одном исключении. В противном случае, когда для всех свидетелей $\rho < 1$, если x бесконечен, эта вероятность равна единице или достоверности, также с одним исключением. Оно имеет место, если величины $\rho_1, \rho_2, \ldots, \rho_x$ монотонно убывают или возрастают и безгранично приближаются к единице.

Пусть для некоторого сомножителя i этого произведения

$$\rho_i = 1 - \frac{4g^2}{(2i-1)^2\pi^2},$$

где π – отношение окружности к своему диаметру и g – некоторая заданная константа, не превышающая единицы, чтобы ни одно ρ_i не стало отрицательным. По известной формуле это произведение будет равно $\cos g$:

$$\sum_i (1 - \frac{4g^2}{(2i-1)^2\pi^2}) = \cos g \text{ и } y_\infty = \frac{q}{q + (1-q)\cos g}$$

и будет весьма отлично от единицы при g отличном от $\pi/2$. Если $g = h\sqrt{-1}$, новая константа h может оказаться меньше или больше единицы и

$$y_\infty = \frac{2q}{2q + (1-q)(e^h + e^{-h})}.$$

Здесь e – основание неперовых логарифмов. Если $h \leq 1$ или хотя бы не очень большое, эта вероятность не будет очень низкой, но легко убедиться, что первое значение y_∞ всегда будет выше, чем вероятность q перед свидетельствами, а второе – всегда ниже.

Эти формулы предполагают, что все свидетельства являются непосредственными, но мы вскоре рассмотрим случай, при котором только одно из них непосредственно, остальные же традиционны.

38. Пусть свидетель вовсе не ограничивается утверждением о том, что что-то истинно или ложно, но указывает на появление происшествия в случае, при котором многие были возможны. Если он ошибается или желает обмануть, то либо оно должно быть лишь одним из не имевших места, либо свидетель вовсе не верит, что оно происходило. Как мы покажем, это обстоятельство влияет на вероятность происшествия, определяемую после свидетельства, вне зависимости от предыдущей вероятности.

Для определённости я предположу, что в урне находится μ шариков с номерами 1, 2, ..., m у $a_1, a_2, \ldots, a_m$ из них, так что

$$\mu = a_1 + a_2 + \ldots + a_m$$

и никаких других номеров нет. При извлечении шарика можно сформулировать m различных гипотез $C_1, C_2, \ldots, C_m$ о его номере. Допустим, что перед свидетельствами их вероятности были равны

$$q_1 = a_1/\mu,\ q_2 = a_2/\mu,\ \ldots,\ q_m = a_m/\mu.$$

Если же свидетель заявляет, что вышел шарик с номером n, в чём и состоит наблюдённое событие, то эти вероятности станут равными $w_1, w_2, \ldots, w_m$. Их требуется определить по правилу § 34.

Наблюдённым событием является свидетельство о появлении номера n. Каждая гипотеза придает этому событию некоторую вероятность p_n, которую нам прежде всего следует определить, так что C_n, q_n, w_n, p_n соответствуют объявленному номеру n. Пусть u будет вероятностью того, что свидетель не ошибается и v – что он не желает обманывать, и, соответственно, $(1 - u)$ и $(1 - v)$ – вероятности противоположных возможностей. В соответствии с

n-й гипотезой свидетель, если он не ошибается и не желает обмануть, сообщает, что вышел шарик с номером n. По правилу § 5 вероятность этого равна uv. Если же он ошибается, то полагает, что был вынут шарик с каким-то другим номером n', а если он в то же время хочет обмануть, то сообщает о выходе шарика с номером, отличном от n', т. е. выбранным из $(m - 1)$ других номеров. Шанс, что он остановится на номере n, будет таким образом равен $1/(m - 1)$, подразумевая, однако, что у него нет никакой большей склонности к какому-либо определённому номеру.

Следовательно, по указанному выше правилу вероятность свидетелю, который ошибается и хочет обмануть, выбрать этот номер равна $(1 - u)(1 - v)/(m - 1)$. Если он ошибся, но не хотел обмануть, или не ошибся и желал обмануть, он не назовёт номер n. В первом случае свидетель хочет сообщить номер, который, как он полагал, действительно вышел, но не был равен n. Во втором случае он знает, что этот номер вышел, но не желает сообщить это. Из всего сказанного по правилу § 10 следует, что полная вероятность того, что гипотеза C_n, если она верна, приводит к наблюдённому событию, равна

$$p_n = uv + \frac{(1-u)(1-v)}{m-1}.$$

В соответствии с гипотезой C_i, $i \neq n$, свидетель, если он не ошибся и не желает обманывать, не назовёт номер n. Если же он не ошибся, но желал обмануть, он знает, что в действительности вышел номер i, но сообщает о выходе одного из остальных $(m - 1)$ номеров, так что шанс выбора n окажется равным $1/(m - 1)$. Поэтому $u(1 - v)/(m - 1)$ будет вероятностью, что свидетель объявит о выходе номера n. Если же он ошибается и не хочет обмануть, эта вероятность станет равной $v(1 - u)/(m - 1)$. Наконец, если свидетель ошибается и желает обмануть, он должен вначале полагать, что вышел один из $(m - 1)$ номеров, о которых он не сообщил, и $1/(m - 1)$ – вероятность, что он считал, что был извлечен какой-то номер n'.

Эта же дробь также выражает вероятность, что он объявит о выходе n, одного из $(m - 1)$ номеров, отличных от n'. Поэтому имеется вероятность $1/(m - 1)^2$ того, что свидетель верит, что вышел n', но объявляет номер n. Шанс этого равен дроби $1/(m - 1)^2$, умноженной на число таких номеров как n', которые, как может полагать свидетель, были извлечены. Подобных номеров только (m – 2), потому что свидетель, который ошибается и желает обмануть, не может полагать, что вышел номер i (что произошло в действительности), или номер n, который он объявил. С другой стороны, вероятность этой двойной ошибки равна $(1 - u)(1 - v)$ и поэтому вероятность, что он объявит о выходе n, будет равна $(1 - u)(1 - v)(m - 2)/(m - 1)^2$. Я объединяю вероятности всех трёх возможных случаев и получаю полную вероятность наблюдённого события в соответствии с одной из (m – 1) гипотез, противоречащих истине

$$p_i = \frac{u(1-v)}{m-1} + \frac{v(1-u)}{m-1} + \frac{(m-2)(1-u)(1-v)}{(m-1)^2}.$$

Эта вероятность связана с вероятностью p_n уравнением

$$p_n + (m - 1)p_i = 1, \qquad (38.1)$$

потому что сумма вероятностей того, что свидетель объявляет о выходе номера n, соответствующая m гипотезам $C_1, C_2, \ldots, C_m$, должна быть равна единице.

Далее, по правилу § 34 мы имеем

$$w_n = \frac{q_n p_n}{q_n p_n + \sum q_i p_i},\; w_i = \frac{q_i p_i}{q_i p_i + \sum q_i p_i},$$

где суммы распространяются на все значения i от 1 до m за исключением n. А поскольку p_i не зависит от i, и сумма значений q_i без включения n равна $(\mu - a_n)/\mu$, выражение для w_n после подстановки значений p_n, q_n, p_i, q_i и умножения числителя и знаменателя на $\mu(m - 1)^2$ становится равным

$$w_n = \frac{\alpha}{\alpha + \beta(\mu - a_n)},\; \alpha = (m - 1)[uv + (1 - u)(1 - v)]a_n,$$

$$\beta = (m - 1)(1 - v)u + (m - 1)(1 - u)v + (m - 2)(1 - u)(1 - v).$$

Такова вероятность, что объявленный свидетелем номер n действительно был извлечён; противоположная вероятность равна $1 - w_n$. В частности, вероятность выхода другого определённого номера i выводится из выражения $1 - w_n$ при умножении его на $q_i p_i/\sum q_i p_i$ или на $a_i/(\mu - a_n)$, так что

$$w_i = \frac{(1-w_n)a_i}{\mu - a_n}.$$

Следует заметить, что для вывода этих результатов мы приняли, что, когда свидетель ошибся или хочет обмануть, объявленный им номер случаен, но не вызван какой-либо причиной. Этого не произойдёт, ни если он желает обмануть, потому что имеет какое-то основание полагать, что один номер будет скорее извлечён, чем другой, или если он ошибается ввиду, скажем, схожести выбранного и объявленного им номера и действительно вышедшего. Эти обстоятельства трудно оценить; мы отвлекаемся от них, но они могут существенно повлиять на вероятность объявленного свидетелем номера.

Вместо пронумерованных шариков урна может содержать такие же количества разноцветных шариков. Если же имеются только белые и чёрные шарики числом a и $\mu - a$, и свидетель сообщил о выходе белого, то в выражении для w_n $m = 2$ и $a_n = a$. Обозначая результат буквой r, мы имеем вероятность, что белый шарик был действительно извлечён:

$$r = \frac{\eta}{\eta + [(1-v)u + (1-u)v](\mu - a)}, \eta = [uv + (1-u)(1-v)]a.$$

Можно представить извлечение белого шарика как частный случай засвидетельствованного истинного или ложного происшествия. Вероятность его истинности будет равна r, а её выражение должно совпадать с данным в § 36. Прежде всего мы имеем

$$p = uv + (1-u)(1-v)$$

для вероятности, что свидетель не обманывает нас. Если он не ошибается и не желает обманывать, или ошибается и хочет обмануть, т. е. в этих двух единственно возможных случаях, извлечения белого и чёрного шариков представляют собой истинность и ложность объявленного события. Свидетель верит противоположному действительности и указывает противное тому, что он полагает. В то же время вероятность обмана равна

$$1 - p = (1-v)u + (1-u)v,$$

что можно вывести из значения p или вычислить сразу же, заметив, что свидетель может обманывать либо когда он не ошибается и хочет этого, либо когда ошибается и не желает обманывать. Кроме того, $q = a/\mu$, $1 - q = (\mu - a)/\mu$ являются вероятностями истинности и ложности события до свидетельства. Эти различные значения приводят к выражению для r из § 36, совпадающему с только что выписанным.

Если каждым из номеров 1, 2, …, m помечен лишь один шарик, то $a_n = 1$ и $\mu = m$, что намного упрощает общее выражение для w_n, которое оказывается равным

$$w_n = uv + \frac{(1-u)(1-v)}{m-1}.$$

Эта вероятность того, что объявленный номер n был действительно извлечён, не отличается здесь от той, которая была выше обозначена через p_n, т. е. от вероятности, что свидетель сообщил о выходе номера n в предположении, что он действительно был извлечён. Она понижается по мере возрастания m и при бесконечном m становится равной вероятности того, что свидетель не ошибся и не желал обмануть.

39. Остаётся рассмотреть общий случай, при котором имеется большое число свидетелей, одни из которых непосредственно знают о событии, другие же знают о нём только по традиции. Но, чтобы изложить это отступление кратко, мы ограничимся решением частного случая подобного рода.

Пусть $T, T_1, T_2, \ldots, T_x$ обозначают $(x + 1)$ свидетелей. Как и в предыдущей задаче, из урны извлекается шарик, о чём свидетель T знает непосредственно, а каждый из остальных узнаёт от предыдущего, что на шарике был указан номер n. Так происходит передача от первого свидетеля T к последнему, T_x, который и сообщает нам о событии по традиционной непрерывной цепи. Свидетель T_x является для нас, следовательно, единственным, которого мы заслушали, и его показание, которое он перенял от T_{x-1}, состоит в том, что номер n был извлечён. Требуется определить вероятность истинности этого сообщения.

Пусть y_x и y'_x – вероятности выхода номеров n и i в соответствии с гипотезами C_n и C_i. Количества шариков с этими номерами обозначим, как всегда, через a_n и a_i, а их общее число – через μ. Тогда априорные шансы их извлечения будут равны a_n/μ и a_i/μ. По правилу § 34 вероятность гипотезы C_n равна

$$w_n = \frac{a_n y_x}{a_n y_x + \sum a_i y'_x}.$$

Сумма распространена на все значения i от 1 до m за исключением n. Сразу же видно, что y'_x не зависит от i, а сумма значений a_i кроме a_n равна $(\mu - a_n)$, так что

$$w_n = \frac{a_n y_x}{a_n y_x + (\mu - a_n) y'_x}.$$

Вероятность w_i любой другой гипотезы C_i можно получить, умножая $(1 - w_n)$ на $a_i/(\mu - a_n)$.

Задача, стало быть, свелась к определению неизвестных y_x и y'_x в функции x. Для её решения я обозначу через k_x вероятность того, что свидетель T_x не обманывает нас, так что $(1 - k_x)$ – вероятность обмана, либо невольного, либо умышленного. Этот свидетель, если он нас не обманывает, объявляет о выходе номера n, если только то же самое заявил свидетель T_{x-1}. Вероятность такого сочетания условий в соответствии с гипотезой C_n равна $k_x y_{x-1}$. Он может объявить о выходе номера n, если обманул нас и в то же время свидетель T_{x-1} сообщил о выходе другого номера. В соответствии с той же гипотезой вероятность этого сочетания равна произведению $(1 - k_x)(1 - y_{x-1})$. Но шанс того, что T_x сообщит о номере n, выбранном из $(m – 1)$ номеров, которые, как он полагает, T_{x-1} не объявил, равна $1/(m – 1)$, так что вероятность объявления n должна быть сведена к $(1 - k_x)(1 - y_{x-1})/(m – 1)$. Наконец, свидетель T_x не сообщит о выходе n либо когда он нас обманывает, а T_{x-1} указывает этот номер, либо когда он не обманывает, а T_{x-1} объявляет какой-то другой номер. Итак, по гипотезам C_n и C_i полная вероятность наблюдённого события равна

$$y_x = k_x y_{x-1} + \frac{(1-k_x)(1-y_{x-1})}{m-1}, \quad y'_x = k_x y'_{x-1} + \frac{(1-k_x)(1-y'_{x-1})}{m-1}.$$

Таким образом, оба неизвестных, y_x и y'_x, включены в одно и то же разностное уравнение первого порядка и отличаются друг от друга лишь на произвольную постоянную. Пусть для y_x она равна c. Тогда общим решением первого уравнения будет

$$y_x = \frac{1}{m} + \frac{c(mk_1 - 1)(mk_2 - 1)...(mk_{x-1} - 1)(mk_x - 1)}{(m-1)^x},$$

и при замене x на $(x - 1)$ окажется, что

$$y_{x-1} = \frac{1}{m} + \frac{c(mk_1 - 1)(mk_2 - 1)...(mk_{x-1} - 1)}{(m-1)^{x-1}},$$
$$1 - y_{x-1} = \frac{m-1}{m} - \frac{c(mk_1 - 1)(mk_2 - 1)...(mk_{x-1} - 1)}{(m-1)^{x-1}}.$$

Эти значения вместе с y_x переводят указанное уравнение в тождество. Чтобы определить c, положим в общем решении $x = 0$ и заметим, что y_x, которое относится к непосредственному свидетелю Т, должно совпадать с p_n § 38. И, при $x = 0$, приняв произведение сомножителей за единицу, мы получим $p_n = (1/m) + c$, $c = (mp_n - 1)/m$. Для некоторого значения x поэтому

$$y_x = \frac{1}{m}[1 + (mp_n - 1)X],\ X = \frac{(mk_1 - 1)(mk_2 - 1)...(mk_x - 1)}{(m-1)^2}.$$

Заметим, что в соответствии с некоторой гипотезой C_i, отличной от C_n, вероятность y'_x для непосредственного свидетеля Т аналогично должна совпадать с p_i § 38 и потому

$$y'_x = \frac{1}{m}[1 + (mp_i - 1)X],$$

не зависия, как и p_i, от i. Я подставляю эти значения в w_n и получаю вероятность, что номер n, объявленный свидетелем T_x, действительно был извлечён:

$$w_n = \frac{[1 + (mp_n - 1))X]a_n}{[1 + (mp_n - 1))X]a_n + [1 + (mp_i - 1))X](\mu - a_n)}.$$

Её и следовало определить. Произведение X можно записать в другом виде и для сокращения письма положить

$$X = h_1 h_2 ... h_x,\quad h_x = k_x - \frac{1 - k_x}{m-1}.$$

Число m всегда более единицы, а k_x – положительная дробь, не превышающая единицу, и поэтому каждый сомножитель произведения X может быть положительным или отрицательным, но не выходить за границы ± 1. Если число x этих сомножителей очень велико, этим произведением можно пренебречь, оно даже равно нулю, если это число бесконечно. Но если сомножители h_1, h_2, … образуют ряд дробей, монотонно стремящихся к единице, указанный случай всё же исключается.

И если в выражении для w_n пренебречь членами, содержащими X, оно становится равным a_n/μ. Отсюда следует, что вообще вероятность события, о котором сообщают очень многие свидетели по традиционной цепи, не отличается существенно от шанса, свойственного ему вне зависимости от свидетельств. Но при очень большом числе непосредственных свидетелей вероятность становится близкой к достоверности, если можно ставить более одного против одного за то, что ни один свидетель нас не обманывает (§ 37).

Если в урне имеется только по одному шарику каждого номера, так что $a_n = 1$ и $\mu = m$, то, ввиду уравнения (38.1),

$$w_n = [1 + (mp_n - 1)X]/m.$$

Эта вероятность таким образом совпадает с y_x, т. е. с вероятностью того, что в соответствии с гипотезой C_n свидетель T_x объявляет о выходе номера n, который действительно был извлечён. Но принимать указанное следствие, $y_x = w_n$, априорно (Laplace 1812/1886, § 44) можно только, если $(\mu - a_n)/a_n = m - 1$.

40. Можно, если угодно, выразить каждую из величин k_1, k_2, … через число m и вероятность u того, что свидетель, к которому они относятся, не ошибается и не желает обманывать. Пусть u_ξ – вероятность того, что свидетель T_ξ из традиционной цепи не ошибся и v_ξ – вероятность того, что он не хочет обманывать. Если эти условия совмещены, то он и не обманывает, как и если ошибается, но хочет обмануть. Но в этом втором случае шанс, что он объявит номер n, равен $1/(m - 1)$. Эти два случая – единственные, при которых свидетель не ошибается, и полное значение k_ξ будет равно

$$k_\xi = u_\xi v_\xi + \frac{(1-u_\xi)(1-v_\xi)}{m-1}.$$

Если принять $u = u_1$ и $v = v_1$, то при $\xi = 0$ оно совпадает с p_n § 38.

Эта величина k_ξ есть вероятность, которую следует приписать свидетелю T_ξ, вероятности которого не ошибиться и не желать обмануть, равны u_ξ и v_ξ или его свидетельству самому по себе, т. е. основанию полагать, что из урны, содержавшей шарики m различных номеров, был извлечён номер n. Если быть уверенным, что этот свидетель не ошибся, но хотел обмануть, то $u_\xi = v_\xi = 0$ и тем не менее $k_\xi = 1/(m-1)$, что достоверно при $m = 2$. Если в этом случае свидетель объявляет номер, в который он не верит, а ошибочно верит, что вышел другой, он неизбежно сообщит истину; при $m = 3$ можно держать пари на равных, что его свидетельство окажется верным. Это можно легко проверить, перенумеровав все возможные варианты, и таким же образом можно проверить значение $1/(m-1)$ вероятности k_ξ при некотором заданном m.

Случай свидетеля, который ошибается и наверняка желает обмануть, нельзя смешивать со случаем разрывной традиционной цепи свидетелей, в которой перед T_ξ отсутствует $T_{\xi-1}$. Ясно, что T_ξ хочет обмануть, поскольку предполагает, что $T_{\xi-1}$ существует, и потому $v_\xi = 0$. Но вероятность не ошибиться у T_ξ вовсе не равна нулю: он ничего не знает о происшедшем событии, и его вероятность назвать верный номер оказывается равной $1/m$. Именно по этой причине таково же значение его свидетельства. И если $u_\xi = 1/m$ и $v_\xi = 0$, то, в соответствии с предыдущей формулой, $k_\xi = 1/m$. Следовательно, $h_\xi = 0$ и $w_n = a_n/\mu$, что и окажется шансом, свойственным выходу n, как это очевидно и должно было быть.

41. Мы здесь имеем в виду завершить сказанное в конце § 7 о приложении правила вероятности причин к склонности нашего разума не сомневаться в том, что некоторые события должны были быть вызваны специальной причиной, не зависящей от случая.

Пусть мы наблюдаем событие, которое само по себе обладает очень низкой вероятностью. Если оно в некотором смысле симметрично, или примечательно в чём-то другом, мы естественно приходим к мысли, что оно не было обусловлено случаем или, общее, какой-то особой причиной, а явилось следствием более мощной причины, как, например, желанием того, кто добивался какой-то особой цели.

Если мы, к примеру[8], увидим на столе расставленные по порядку 26 напечатанных букв алфавита, a, b, c, …, x, y, z, у нас не будет ни малейшего сомнения в том, что кто-то хотел их так расположить. И всё же само по себе это расположение не менее правдоподобно, чем любое другое, не обладающее ничем примечательным, которое мы по этой причине не колеблясь полностью припишем случаю. Если те же 26 букв последовательно и случайно извлечены из урны, шанс их выхода либо в указанном порядке, либо в каком-то другом произвольно выбранном порядке, например, b, p, w, …, q, a, t, окажется тем же самым и столь же малым, но не меньшим.

Аналогично, если урна содержит равное число белых и чёрных шариков, и мы должны последовательно извлечь с возвратом 30 из них, то вероятность, что все они окажутся белыми, будет равна $1/2^{30} \approx 10^{-9}$. Но вероятность любого иного результата окажется не выше и не ниже, а такой же. Тем не менее, увидев последовательный выход 30 белых шариков, мы не сможем поверить, что это произошло случайно. В то же время мы без труда припишем случаю ничем не примечательное появление стольких же шариков.

То, что мы назвали *случаем* (§ 27), так же легко, так сказать, производит и события, которые мы посчитаем примечательными, и обычные. Если равновозможных событий очень много, то первые намного более редки и по этой причине они и поражают наше воображение и заставляют нас искать специальную причину их возникновения. И её существование весьма вероятно, что, однако, не вызвано их редкостью. Она основана на другом принципе, к которому мы применим доказанные ранее правила.

42. Обозначим возможные примечательные события E_1, E_2, …, а обычные – F_1, F_2, … Так, извлечения из урны, содержащей равное число белых и чёрных шариков, 30 шариков одного и того же цвета, или 15 одного цвета, а вслед за этим – 15 другого цвета, будут событиями E_1, E_2. В случае примерно 30 напечатанных букв этими событиями будут их расположение либо в алфавитном или обратном порядке, либо составление из них фразы французского или иного языка.

Пусть первых будет m, а вторых – n, и пусть все они равно возможны, если обусловлены только случаем. Вероятность каждого тогда окажется равной $p = 1/(m + n)$. Но условия изменятся, если эти события были вызваны особой причиной C, не зависящей от вероятности p; скажем для определённости, вызваны чьим-то желанием и его выбором. Примем, что этот выбор определяется различными обстоятельствами, которые устанавливают, что некоторые возможные события станут примечательными. Итак, существуют вероятности p_1, p_2, … того, что выбор указанного лица приведёт к E_1, E_2, … но не будет никакой вероятности появления событий F_1, F_2, … Тогда

$$p_1 + p_2 + \ldots = 1,$$

а если они совпадают, то каждая равна $1/m$ и очень высока и сама по себе, и относительно p, если общее число возможных случаев очень велико само по себе и относительно m.

Вообще говоря, эти вероятности могут весьма различаться, но мы не имеем никакой возможности определить их. Впрочем, нам достаточно знать, что они намного выше, чем p; иначе не может быть, потому что вероятность p очень низка, т. е. число $m + n$ особо велико, как в примерах ниже. Таков принцип, от которого мы исходим при определении вероятности причины С после наблюдения одного из событий E_1, E_2, …, F_1, F_2, … или по меньшей мере при установлении, что она очень высока, если наблюдённое событие относится к примечательным.

Назовём это событие E_1; можно сформулировать две гипотезы, а именно, что оно было вызвано либо причиной С, либо случаем. Если первая гипотеза достоверна, то p_1 будет вероятностью появления E_1, если же вторая, то её вероятность окажется равной p. Пусть r будет вероятностью первой гипотезы после наблюдения, до которого обе они были равновероятны. Тогда по правилу § 28 $r = p_1/(p_1 + p)$. Для того, чтобы r очень мало отличалось от единицы или достоверности, достаточно, чтобы p_1 было очень велико сравнительно с очень небольшим шансом p. В одном из предыдущих примеров число возможных событий превышало 10^9, а p было меньше $1/10^9$. Если предположить, что число событий, достаточно примечательных, чтобы выбирать одно из них [остановиться на одном из них], равно 1000, и принять 1/1000 за значение p_1, то r будет отличаться от единицы меньше, чем на одну миллионную, – и ещё намного меньше, если, как можно полагать, p_1 превышает одну тысячную. Если одно из таких примечательных событий произошло, как, например, выход 30 шариков одного и того же цвета из урны, содержащей равное число шариков двух различных цветов, то это событие следует без всяких сомнений, как это, естественно, и делается, приписать желанию кого-то или какой-то совсем иной специальной причине, но не считать его простым проявлением случая.

Однако, вероятность r причины С намного снизится, если перед наблюдением её существование или отсутствие не были равновозможными, как предполагалось в предыдущей формуле, и что более вероятным было её отсутствие. Так и будет в следующем примере, в котором перед тиражом было принято достаточно предосторожностей, чтобы исключить влияние любого пожелания на извлечение шариков. Учитывая это обстоятельство, правило § 34 ощутимо уменьшает значение *r*. Эта вероятность также возрастёт или уменьшится, иногда в громадной степени, когда события E_1, E_2, …, F_1, F_2, … никак не равно возможны. Именно, она возрастёт, если шанс, свойственный каждому событию, меньше у примечательных, чем у обычных из них, и уменьшится в противном случае.

Согласованность, которую мы наблюдаем в природе, без сомнения не вызвана случаем. Внимательное и длительное исследование очень многих явлений обнаружило физические причины их возникновения, если не с полной достоверностью, то с вероятностью очень близкой к единице. Можно считать эти явления примечательными вещами E_1, E_2, …; их вероятности настолько высоки сами по себе, что вмешательство причины, подобной той, которую мы обозначили C, весьма маловероятно, и иногда даже бесполезно отыскивать её. По поводу физических явлений, причины которых ещё не изучены, разумно приписывать им аналогичные известным и считать, что они подчинены тем же законам. С прогрессом наук их число, впрочем, уменьшается со дня на день[9]. Сегодня, к примеру, мы знаем причину появления молнии и понимаем, как планеты удерживаются на своих орбитах, что не было известно нашим предшественникам, а будущие поколения узнают ныне неизвестные причины других явлений.

43. Если число различных причин, которые могут быть приписаны наблюдённому событию E, бесконечно, их вероятности и до, и после появления E, становятся бесконечно низкими, и суммы в формулах §§ 32 и 34 переходят в определённые интегралы. Чтобы осуществить этот переход, предположим, что E состоит в извлечении белого шарика из урны, содержащей бесконечное количество белых и чёрных шариков. О неизвестном соотношении *x* числа белых к общему числу шариков можно сформулировать бесконечно много гипотез и принять их за такое же количество различных взаимоисключающих причин появления E, так что *x* сможет принимать все значения, возрастающие бесконечно малыми приращениями, от бесконечно малых (единственный белый шарик в урне) до единицы (только белые шарики находятся в урне).

Пусть X будет вероятностью, что это соотношение, если его значение x достоверно, приведёт к появлению Е. В любом случае X окажется известной функцией x, и, если рассматривать это значение [того же x] как возможную причину наступления Е, то речь пойдёт об определении бесконечно малой вероятности x, когда либо все причины до наблюдения равновероятны, либо их априорные шансы различны.

В первом случае требуемая вероятность выводится из w_m (§ 28), полагая, что m бесконечно и подставляя вместо p_1, p_2, … значения x, соответствующие X. Используя символ суммы как в § 32 и обозначая вероятность x через w, мы получаем $w = X:\sum X$, и в соответствии с фундаментальной теоремой определённых интегралов

$$\sum Xdx = \int_0^1 Xdx.$$

Следовательно, если считать dx постоянным и умножить на него числитель и знаменатель предыдущей дроби w, то

$$w = Xdx \div \int_0^1 Xdx.$$

Пусть X' будет вероятностью, соответствующей x для будущего события Е′, которое зависит от тех же причин, что и Е, и w' – полной вероятностью появления Е′. Тогда, подставляя вместо w его предшествующее значение и заменяя сумму интегралами, мы получим по правилу § 30

$$w' = \sum X'w = \int_0^1 XX'dx \div \int_0^1 Xdx.$$

Если же перед наблюдением Е различные значения x не равновероятны, то мы обозначим через Ydx бесконечно малую вероятность шанса x этого события и заменяем q_n в формулах § 34 на это Ydx. Так мы получим вероятность этого шанса после появления Е

$$w = XYdx \div \int_0^1 XYdx,$$

а вероятность появления будущего события E′ окажется равной

$$w' = \int_0^1 XX'Ydx \div \int_0^1 XYdx.$$

44. Если заранее известно, что x не может принимать значения от 0 до 1, а заключён в заданные границы, они и должны быть приняты в качестве пределов определённых интегралов. Если желательно сохранить прежние пределы, следует предположить, что Y – разрывная функция x, равная нулю вне заданных пределов этой переменной. В любом случае, принимай x все значения от 0 до 1, или будь он более ограничен, если обозначить через λ вероятность, что его неизвестное значение после появления E заключено в более узкие границы α и β, то она окажется суммой значений w, соответствующих этим последним границам

$$\lambda = \int_\alpha^\beta XYdx \div \int_0^1 XYdx.$$

При приближённых вычислениях эта формула применима, когда число причин, которые могут быть приписаны событию E, не бесконечно, а лишь весьма значительно. Пусть, например, E состоит в последовательном извлечении с возвратом n белых шариков из урны, содержащей очень большое число белых и чёрных шариков. Вероятность X события E, соответствующая соотношению x числа белых шариков к общему числу шариков, окажется равной этому соотношению в степени n. Если отыскивается вероятность того, чтобы белых шариков в урне было больше, чем чёрных, в выражении для λ следует принять $\alpha = 1/2$ и $\beta = 1$. А если кроме того до тиражей все возможные значения x равновероятны, то Y не будет изменяться с x и потому исчезнет из этого выражения, так что

$$X = x^n, \int_0^1 Xdx = \frac{1}{n+1}, \int_{1/2}^1 Xdx = \frac{1}{n+1}(1 - \frac{1}{2^{n+1}})$$

и поэтому $\lambda = (1 - 1/2^{n+1})$.

Это окажется тем точнее, чем больше чёрных или белых шариков находится в урне. До тиражей можно держать пари на равных, что белых шариков больше, чем чёрных. Достаточно, однако, извлечь белый шарик, чтобы оказалось $\lambda = 3/4$, а если последовательно извлечь некоторое небольшое число таких же шариков, вероятность λ того, что белых шариков в урне больше, чем чёрных, окажется почти достоверностью.

45. Как я уже сказал (§ 30), E и E′ можно считать событиями, составленными из одного и того же простого события G, и связанными друг с другом этой своей общей зависимостью от G[10]. Шанс G неизвестен; до появления E вероятность его значения x равнялась Ydx, а после этого события равна w. А так как этот шанс наверняка заключён между 0 и 1, сумма значений соответствующих Ydx равна единице, как и было для суммы значений w.

Данная функция Y от x, будь она непрерывна или разрывна, должна удовлетворять условию $\int Ydx = 1$. По правилу математического ожидания (§ 23), применяя его к шансу G, следует за его значение принять до наблюдения E сумму всех его возможных значений, умноженных на их соответствующие вероятности, т. е. сумму всех произведений $xYdx$ от $x = 0$ до $x = 1$. Обозначим этот шанс через γ; точнее, γ будет то, что следует принять за его неизвестное значение до наблюдения E:

$$\gamma = \int_0^1 xYdx.$$

Если считать x и Y абсциссой и ординатой плоской кривой, так что её полная площадь $\int Ydx$ равна 1[11], то γ будет абсциссой центра тяжести этой площади. Приняв это значение γ за шанс G, можно держать пари о появлении этого события в первом же испытании, но не о многих его последовательных наступлениях, потому что, появится ли G при первом испытании или нет, вероятность его дальнейшего наступления соответственно либо возрастёт, либо уменьшится.

Если, например, до наблюдения все значения x были равновероятны, Y станет независимым от x и, в соответствии с двумя предыдущими уравнениями,

$$Y = 1,\ \gamma = 1/2.$$

По существу, у нас нет никаких оснований полагать, что при первом испытании появится скорее G, а не противоположное событие. Но если заменить E и E′ этим простым событием, то

$$X = x,\ X' = x \text{ и } w' = \int_0^1 XX'dx \div \int_0^1 Xdx \ = \ \frac{2}{3}, -$$

вероятность, что G, появившееся при первом испытании, наступит и во втором, поскольку она повысилась на 1/6 [с 1/2 до 2/3].

Она понизится на ту же величину, если в первом испытании появилось противоположное событие. Действительно, если принять это последнее за E, и, как всегда, G – за E′, т. е. положить

$$X = 1 - x,\ X' = x, \text{ то } w' = \int_0^1 x(1-x)dx \div \int_0^1 (1-x)dx \ = \ \frac{1}{3}$$

будет вероятностью того, что G, не наступившее при первом испытании, появится во втором. До испытаний вероятность, что G наступит дважды подряд, будет по правилу § 9 произведением вероятностей 1/2 и 2/3, т. е. 1/3 вместо 1/4, как было бы, будь вероятность G в каждом испытании равна 1/2. Вероятность наступления одного и того же события в двух первых испытаниях, т. е. либо только G, либо, что равновероятно, только противоположного события, равна 2/3. Сравнивая 2/3, или (1/2)(1 + 1/3), с вероятностью $(1/2)(1 + \delta^2)$ для повторения событий (§ 26), мы получаем $\delta = 1/\sqrt{3}$. Если у нас вначале нет никаких сведений о шансе G, мы можем с одним и тем же основанием предположить, что x принимает все возможные значения[12].

Мы сейчас определим вероятность совпадения исходов, если заранее известно, что возможные значения x не равновероятны, а весьма вероятно чуть отличаются друг от друга на известную или неизвестную дробь.

46. Простое событие, шанс которого неизвестен, мы как всегда обозначаем через G. Назовём противоположное событие H, так что его шанс равен единице без шанса G, и предположим *первое*, что наблюдённое событие E состоит в наступлении в некотором порядке G и H m и n раз. И *второе*, что будущее событие E′ состоит в наступлении тех же событий, снова в некотором порядке, m' и n' раз.

При значении x шанса G и $1 - x$ – шанса H вероятности X и X' событий E и E′ будут (§ 14)

$$X = Kx^m(1 - x)^n,\ X' = K'x^{m'}\ (1 - x)^{n'},$$

где K и K' – числа, не зависящие от x. Тогда вероятность E′ после наступления E будет равна

$$w' = K'\int_0^1 Yx^{m+m'}(1-x)^{n+n'}dx \div \int_0^1 Yx^m(1-x)^n dx.$$

Число K исчезло из этой формулы, а $K' = C_{m'+n'}^{m'}$ Если E′ состоит в появлении G и H в определённом порядке m' и n' раз, то следует считать $K' = 1$.

Далее, если до наблюдения E не было ни малейшего основания полагать какое-либо значение x более вероятным, чем другое, то $Y = 1$ и при интегрировании по частям

$$\int_0^1 x^m(1-x)^n dx = \frac{n!}{(m+1)(m+2)...(m+n)(m+n+1)} =$$
$$m!n!/(m + n + 1)!.$$

Так же само

$$\int_0^1 x^{m+m'}(1-x)^{n+n'}dx = \frac{(m+m')!(n+n')!}{(m+m'+n+n'+1)!}$$

и, ввиду $K' = C_{m'+n'}^{m'}$, в нашем случае

$$w' = \frac{(m'+n')!(m+m')!(n+n')!(m+n+1)!}{m'!n'!m!n!(m+m'+n+n'+1)!}.$$

Чтобы эта формула включала случай, при котором одно из чисел m, n, m', n' равнялось нулю, следует считать, что $0! = 1$. Тогда при $n = n' = 0$

$$w' = \frac{m+1}{m+m'+1}.$$

Эта формула выражает вероятность последовательного появления G μ раз после последовательного появления противоположного события H m раз. При $m' = 1$, $n' = 0$, $Y = 1$ и

$$w' = \frac{m+1}{m+n+2},$$

а при $m' = 0$, $n' = 1$

$$w' = \frac{n+1}{m+n+2}.$$

Сумма этих двух дробей равна единице, что и должно было быть, потому что первая выражает вероятность, что G появится в следующем после $(m + n)$ испытаний, а вторая дробь, – что не появится. Первая больше или меньше второй в зависимости от $m > n$ и $m < n$, или иначе, наступало ли G чаще, чем противоположное событие H, или реже. Дроби равны друг другу и равны половине, как и до испытаний, если оба указанные события произошли одно и то же число раз.

Но, вообще говоря, положение изменится, если заранее будет известно, либо по самой сути события G, либо по результатам испытаний, проведенных до наступления события E, что значения неизвестного шанса G не равновероятны, так что $Y \neq 1$. Дробь γ из § 45, которую следовало принять за шанс G перед $(m' + n')$ новыми испытаниями, в этом случае вовсе не равна 1/2. При следующем испытании вероятность G может быть ниже, хоть это событие появлялось чаще, чем противоположное событие H, или выше, хотя бы G появлялось реже, чем H. Это будет показано в следующем примере.

47. Я предположу, что заранее весьма вероятно, что шанс G очень мало отличается в ту или иную сторону от некоторой определённой дроби r, так что

$$x = r + z.$$

Величина Y будет функцией z, которая принимает ощутимые значения лишь при очень небольших положительных или отрицательных значениях этой переменной. Плоская кривая, для которой x и Y являются текущими координатами, будет заметно отходить от оси абсцисс $0x$ лишь на очень небольшом интервале по обе стороны от ординаты $x = r$. Центр тяжести площади этой кривой поэтому окажется в указанном интервале в точке с абсциссой, очень мало отличающейся от r. Если пренебречь этой разностью, то r будет значением γ из § 45.

Заметив это, укажем, что пределами интегралов будут $z = -r$ и $1 - r$, что соответствует значениям $x = 0$ и 1. Поэтому в первом выражении для w' из § 46 $m' = 1$, $n' = 0$, $dx = dz$ и

$$w' = \int_{-r}^{1-r} Yx^{m+1}(1-x)^n dz \div \int_{-r}^{1-r} Yx^m (1-x)^n dz$$

оказывается вероятностью, что G появится снова после наступления m раз, а H – n раз в $(m + n)$ испытаниях. Однако, по сути сомножителя Y можно, если угодно, ограничить пределы интегралов весьма малыми значениями z. Если разложить остальные сомножители в ряды по степеням z, то, как правило, эти ряды будут быстро сходиться. Исключением окажется только случай, при котором r или $1 - r$ также будут очень небольшими дробями. Во всех остальных случаях достаточно сохранять первые члены разложений, и, пренебрегая квадратом z [?], вывести

$$x^m(1-x)^n = r^m(1-r)^n + [mr^{m-1}(1-r)^n - nr^m(1-r)^{n-1}]z +$$
$$½[m(m-1)r^{m-2}(1-r)^n - 2mnr^{m-1}(1-r)^{n-1} +$$
$$n(n-1)r^m(1-r)^{n-2}]z^2,$$

откуда при замене m на $(m + 1)$ определяется значение $x^{m+1}(1-x)^n$.

Я подставляю эти значения $x^m(1-x)^n$ и $x^{m+1}(1-x)^n$ в выражение для w'. Если заметить, что

$$\int_{-r}^{1-r} Ydz = 1, \quad \int_{-r}^{1-r} Yzdz = 0,$$

принять для сокращения письма

$$\int_{-r}^{1-r} Yz^2dz = h$$

и пренебречь лишь весьма малой дробью h^2, то

$$w' = r + \left(\frac{m}{r} - \frac{n}{1-r}\right)h.$$

Это доказывает, что вероятность w' появления G после $m + n$ испытаний будет больше, если $[m/r < n/(1 - r)]$, или меньше, если $[m/r > n/(1 - r)]$, дробей r или γ, которые следует принять за шанс G перед испытаниями.

Если достоверно, что r было шансом G, а m и n были большими числами, то G и H будут весьма вероятно появляться пропорционально своим шансам r и $(1 - r)$, а при совпадении m/r и $n/(1 - r)$ вероятность w' станет равной шансу r, как и должно было быть.

48. Если $m = 1$ и $n = 0$, то $w' = r + h/r$ будет вероятностью того, что G, наступившее при первом испытании, появится и во втором. А так как r есть вероятность его наступления в первом испытании, $rw' = r^2 + h$ выразит вероятность его повторного появления. Принимая $(1 - r)$ вместо r, мы получим $(1 - r)^2 + h$ для вероятности повторения противоположного события. Если добавить $r^2 + h$, сумма

$$1 - 2r + 2r^2 + 2h$$

окажется вероятностью повторения одинаковых результатов в этих двух испытаниях.

Пусть в выражениях для w' из § 47 $m = 0$ и $n = 1$, тогда $w' = r - h/(1 - r)$ окажется вероятностью того, что G, не наступившее в первом испытании, появится во втором, а произведение $r(1 - r) - h$ будет вероятностью появления G и H. Удвоив его, мы получим вероятность повторного появления противоположных событий, что можно было бы вывести, вычитая вероятность повторения совпадающих событий из единицы. Разность вероятностей этих двух исходов, последняя из которых выше первой, равна

$$(1 - 2r)^2 + 4h.$$

Она возросла потому, что r не является точным шансом G, и мы только знаем, что этот шанс очень мало отличен от r. Если бы нам было также известно, что r равнялось 1/2, всё же мы могли бы с выгодой держать пари на равных на то, что исходы будут совпадать. Это и имеет место при игре в орлянку перед первым броском монеты. Равенство шансов обоих исходов физически невозможно, и из метода чеканки монет следует, что с высокой вероятностью шанс каждого очень немного отличается от 1/2.

49. Я уже здесь укажу доказываемую в следующей главе теорему, которая послужит для определения шанса события по опыту с очень высокой вероятностью, не достоверно и строго, но с весьма хорошим приближением.

Пусть g – известный или неизвестный шанс события G, т. е. отношение числа равновозможных благоприятных ему случаев ко всем возможным и также равновозможным случаям. Предположим, что произведено μ испытаний, в течение которых этот шанс, свойственный G и отличный от вероятности наступления этого события (§ 1), остаётся постоянным. Отношение r количества появления G к μ может сильно изменяться с μ и отличаться в одну или другую сторону от g до тех пор, пока μ не слишком велико. Но когда μ станет большим, разность $r - g$ без учёта знака будет всё более уменьшаться по мере дальнейшего возрастания μ и, если μ сможет стать бесконечным, окажется строго равной нулю. Обозначив через ε произвольно малую дробь, можно всегда установить такое большое μ, при котором вероятность того, что $r - g$ окажется меньше, чем ε, сколь угодно приблизится к достоверности. В дальнейшем мы укажем выражение для вероятности неравенства $r - g < \varepsilon$ в функции μ и ε.

Итак, пусть урна содержит a белых и b чёрных шариков. Из неё последовательно извлекают с возвращением очень большое число μ шариков, из которых α белых и β чёрных. Тогда тем точнее и с тем более высокой вероятностью, чем больше $\mu = \alpha + \beta$,

$$\frac{\alpha}{\alpha+\beta}=\frac{a}{a+b},\ \frac{\beta}{\alpha+\beta}=\frac{b}{a+b},\ \frac{\alpha}{\beta}=\frac{a}{b}.$$

Обратно, если отношение чисел белых и чёрных шариков в урне неизвестно, после очень большого числа испытаний, в течение которых это отношение не изменилось, можно с очень высокой вероятностью считать приближёнными значениями этого неизвестного отношения и неизвестного шанса извлечения белого шарика величины α/β и α/(α + β), каково бы ни было число шариков в урне.

Всё же следует заметить, что если a очень невелико по сравнению с b, то α окажется очень малым сравнительно с β и соответственно в противном случае. Отношение одной из дробей α/β и a/b к другой может намного отличаться от единицы, по крайней мере пока серия испытаний не продвинется исключительно далеко. Если известный или неизвестный шанс наступления белого шарика очень слаб, приближённое равенство $\alpha/\beta = a/b$ лишь укажет, что обе эти дроби очень малы.

Указанное правило, относится равным образом к шансам различных взаимоисключающих причин, которые могут быть приписаны событию E, наблюдаемому очень большое число раз. Если γ – известный или неизвестный шанс одной из таких причин C, дробь $\gamma/(1-\gamma)$ очень близко и с высокой вероятностью окажется отношением числа появлений E, вызванных причиной C, к числу его появлений, порождённых всеми остальными причинами. Так можно определить это отношение, если шанс γ был известен заранее или его значение было определено по опыту.

Пусть E является, к примеру, выходом белого шарика из урн A и B, содержащих a белых и a' чёрных и b и b' белых и чёрных шариков соответственно. Значения шансов γ и $(1-\gamma)$ урнам A и B быть причинами E по правилу § 28 равны

$$\gamma = \frac{a(b+b')}{a(b+b')+b(a+a')},\ 1-\gamma = \frac{b(a+a')}{a(b+b')+b(a+a')}.$$

После выхода очень большого числа μ белых шариков из одной или другой урны с возвратом вышедших шариков того или иного цвета в исходные урны отношение чисел вышедших из A и B будет весьма вероятно очень мало отличаться от $\gamma/(1-\gamma)$, так что можно принять, что

$$\rho = \frac{\gamma}{1-\gamma} = \frac{a(b+b')}{b(a+a')}.$$

Если $a + a' = b + b'$, то $\rho = a/b$. В этом случае можно объединить все шарики в одной урне D (§ 10) не изменяя соотношения тех белых, вышедших из A и B, которые будут извлечены из D. При очень большом числе белых шариков первое из этих чисел будет относиться ко второму почти как a/b. Это можно проверить, если как-то пометить различным образом шарики, бывшие в урнах A и B.

50. В трудах Бюффона (1777, § 18) можно найти численные результаты опыта игры в орлянку, которые послужат примером и подтвердят предыдущие правила. В этой игре шансы исходов броска монеты зависят от её недостаточно известного нам физического состава, и даже в противном случае установление этих шансов явилось бы задачей механики, которую никто не мог бы решить. Поэтому их приближённые значения должны быть установлены опытом для каждой монеты.

Если после очень большого числа испытаний μ орёл появился m раз, отношение m/μ должно быть принято за его шанс. Это же отношение является и его вероятностью, т. е. основанием полагать, что он появится при новом испытании той же монеты, а после завершения их серии можно будет справедливо ставить m против $\mu - m$ на появление орла. Также, исходя из этой вероятности m/μ простого события, следует вычислять вероятности составных событий, по крайней мере, если они не слишком низки по своей сути.

Предположим теперь, что имеется очень большое число m серий испытаний, каждая из которых продолжается, как это и было в эксперименте Бюффона, до появления орла. Пусть это происходило при первом, втором, … броске монеты a_1, a_2, … раз. Общее число бросков или испытаний μ и появлений орла m равно

$$\mu = a_1 + 2a_2 + 3a_3 + \dots,\ m = a_1 + a_2 + \dots$$

Шанс появления орла будет равен $p = m/\mu$ с теми лучшими приближением и точностью, чем больше возрастёт μ. Вероятности появления орла при первом броске монеты; только при втором; третьем, … равны p, $p(1 - p)$, $p(1 - p)^2$, … В соответствии с гипотезой количества появлений этих событий в m сериях испытаний равны a_1, a_2, … и поэтому, если m – очень большое число, а вероятности не становятся очень малыми дробями, почти точно

$$p = a_1/m,\ p(1 - p) = a_2/m,\ p(1 - p)^2 = a_3/m,\ \dots$$

Разделив каждое из этих уравнений на предыдущее, мы получим различные значения $1 - p$, и, следовательно,

$$p = a_1/m = 1 - a_2/a_1 = 1 - a_3/a_2 \dots$$

Эти значения p, или по крайней мере некоторые из его первых значений, будут тем меньше отличаться друг от друга и от отношения m/μ, чем большими будут числа m и μ. Чтобы они стали наверняка тождественными, требуется бесконечность указанных чисел. Полагая p равным среднему из этих весьма мало различающихся дробей, или значению m/μ, мы имеем для вычисленных значений $a_1, a_2, \ldots$

$$a_1 = mp,\ a_2 = mp(1-p),\ a_3 = mp(1-p)^2, \ldots$$

Они, или по крайней мере первые члены этой убывающей геометрической прогрессии, должны очень немного отличаться от их наблюдённых значений. В эксперименте Бюффона $m = 2048$ и в соответствии с его отчётом

$$a_1 = 1061,\ a_2 = 494,\ a_3 = 232,\ a_4 = 137,\ a_5 = 56,$$
$$a_6 = 29,\ a_7 = 25,\ a_8 = 8,\ a_9 = 6.$$

Чисел $a_{10}, a_{11}, \ldots$ нет, потому что число m не было достаточно большим, и орёл неизменно появлялся в одном из предыдущих бросков.

Число m равно сумме значений $a_1, a_2, \ldots$, а $\mu = 4040$ и $p = m/\mu = 0.50693$. Исходя из этого значения[13] и пренебрегая дробями, мы получаем

$$a_1 = 1038,\ a_2 = 512,\ a_3 = 252,\ a_4 = 124,\ a_5 = 61,$$
$$a_6 = 30,\ a_7 = 15,\ a_8 = 7,\ a_9 = 4,\ a_{10} = 1,$$

причём числа $a_{11}, a_{12}, \ldots$ оказались меньшими единицы.

Сравнивая эти вычисленные числа с измеренными, мы видим, что их первые члены мало отличаются друг от друга. Расхождения между последующими членами более значительны; например, a_7 составляет лишь 3/5 от наблюдённой величины, однако оно соответствует событию, вероятность которого ниже 1/100. Останавливаясь на трёх первых наблюдённых членах серии, можно вывести

$$p = a_1/m = 0.51806,\ p = 1 - a_2/a_1 = 0.53441,$$
$$p = 1 - a_3/a_2 = 0.53033.$$

Эти величины очень немного отличаются друг от друга, а их среднее, $p = 0.52760$, едва отличается на 0.02 от значения m/μ для всех испытаний в целом. Я выбрал этот эксперимент ввиду репутации автора, а также и потому, что он стал достоверным, поскольку был включён в его труд. Каждый может выполнить много других экспериментов того же вида, подбрасывая либо монету, либо шестигранную игральную кость. В последнем случае при обширной серии наблюдений число появлений каждой из них окажется примерно равным 1/6 количества испытаний, если только кость не является ни неправильной, ни плохо изготовленной.

51. Теорема, на которой основано предыдущее правило, обязана своим возникновением Якобу Бернулли, 20 лет размышлявшему о её доказательстве, которое следует из биномиальной формулы при помощи следующих предложений.

Пусть в каждом испытании p и q – данные шансы противоположных событий E и F, а g, h, k – такие целые числа, что

$$p = g/k,\ q = h/k,\ g + h = k,\ p + q = 1.$$

Обозначим через m, n, μ другие целые числа, связанные с предыдущими уравнениями

$$m = gk,\ n = hk,\ \mu = m + n = (g + h)k,$$

так что p относится к q как m к n, которые без изменения их отношения могут быть произвольно увеличены соответственно увеличенным g, h и k. Поэтому

[1] Наибольшим членом разложения $(p + q)^\mu$ является тот, который содержит $p^m q^n$. Поскольку он выражает вероятность того, что E и F наступят m и n раз (§ 14), это составное событие, т. е. появление E и F пропорциональное их шансам, является наиболее вероятным из всех составных событий, которые могут происходить в некотором числе μ испытаниях.

[2] Если μ очень велико, отношение наибольшего члена разложения $(p + q)^\mu$ к сумме всех членов или к единице будет очень малой дробью, к тому же безгранично убывающей с дальнейшим ростом μ. Следовательно, в длинной серии испытаний это отношение окажется очень малым и будет всё более убывать с дальнейшим удлинением серии испытаний.

[3] Но если в разложении $(p + q)^\mu$ рассматривать его наибольший член и l членов до и после него и обозначить через λ сумму этих $2l + 1$ последовательных членов, всегда можно будет без изменения p и q принять такое достаточно большое μ, чтобы дробь λ отличалась от единицы сколь угодно мало. И если μ будет и далее возрастать, λ станет всё ближе приближаться к единице.

Отсюда следует, что в длинной серии испытаний всегда существует высокая вероятность λ[14] того, что количество наступлений события E будет заключаться в границах $(m \pm l)$, а событий F – в границах $(n \mp l)$. Не изменяя интервала $2l$ можно поэтому выбрать такое большое μ, что вероятность λ также приблизится сколь угодно близко к единице. С учётом предыдущих уравнений, взяв эти отношения и обозначив

$$l/\mu = \delta,\ p \pm \delta = p',\ q \mp \delta = q',$$

мы замечаем, что они равны p' и q'. А так как дробь δ безгранично убывает с ростом μ, эти отношения l/μ, изменяющиеся вместе с μ, также с очень высокой вероятностью безгранично приближаются к шансам p и q событий E и F. В этом и состоит прекрасная теорема Якоба Бернулли[15].

Для доказательства этих свойств членов разложения $(p + q)^\mu$ мы отсылаем читателей к четвёртой части *Искусства предположений* и первому разделу книги Lacroix (1816). Что касается самой теоремы, о которой см. в следующей главе, она основана на интегральном исчислении. Нельзя, однако, забывать, что она существенно предполагает, что шансы простых событий E и F в течение испытаний остаются неизменными. Но в приложениях исчисления вероятностей и к различным физическим явлениям, и к моральным обстоятельствам они чаще всего изменяются от одного испытания к другому, и также чаще всего вполне нерегулярно. Теорема Якоба Бернулли не подходит для подобных случаев, но существуют общие предложения, которые учитывают какие угодно изменения шансов, и служат основанием наиболее важных приложений теории вероятностей. Они также доказываются в следующих главах. Пока же мы опишем и выведем *закон больших чисел*, который был упомянут в Предисловии, в качестве общего факта, происходящего при наблюдениях любой природы.

52. Предполагая очень большое число μ последовательных испытаний, обозначим шансы события E какой-либо природы через $p_1, p_2, \ldots, p_\mu$, а их среднее через $\overline{p}$, а $(1 - p_1), (1 - p_2), \ldots, (1 - p_\mu)$ будут шансами противоположного события F с их средним $\overline{q}$, $\overline{p} + \overline{q} = 1$. Одно из общих предложений, которые мы имеем в виду рассмотреть, состоит в том, что, если обозначить через m и n количества появлений E и F в серии испытаний, отношения m и n к $\mu = m + n$ будут с очень высокой вероятностью почти точно равны $\overline{p}$ и $\overline{q}$, и обратно, $\overline{p}$ и $\overline{q}$ окажутся приближёнными значениями m/μ и n/μ.

[1] Если эти отношения выведены из длинной серии испытаний, шансы $\overline{p}$ и $\overline{q}$ событий E и F станут известны, а по правилу § 49 определятся сами шансы p и q, будь они постоянны. Но чтобы эти приближённые значения $\overline{p}$ и $\overline{q}$ могли послужить, также приближённо, для оценки количества появлений E и F в новой длинной серии испытаний, требуется, чтобы средние шансы E и F достоверно или по крайней мере весьма вероятно оставались точно теми же, или почти теми же. И это имеет место в соответствии с другим общим предложением, которое мы сейчас выскажем.

[2] Пусть по своей природе события E и F, которые появляются в каждом испытании, могут быть вызваны одной из взаимоисключающих причин $C_1, C_2, \ldots, C_\nu$, которые я вначале предположу равновозможными. Я обозначу через c_i шанс того, что какая-то причина C_i привела к появлению E. Если существует лишь одна возможная причина, шанс E будет с необходимостью одним и тем же во всех испытаниях, однако в соответствии с нашей гипотезой при каждом испытании возможны его ν равновероятных значений, и он поэтому изменяется от одного испытания к другому. Средний шанс события E в очень большом числе прошедших или будущих испытаний весьма вероятно окажется почти равным γ, чьё значение не зависит от этого числа. Тогда средний шанс события E можно считать неизменным в двух или большем числе очень длинных сериях испытаний.

Объединяя это второе общее предложение с первым, мы заключаем, что если m – число появлений события E в очень большом числе μ прошедших или будущих испытаний, а m' и μ' – соответствующие числа для второй очень длинной серии, то с высокой вероятностью почти точно $m/\mu = m'/\mu'$. Если μ и μ' могут стать бесконечными, равенство окажется точным, и оба отношения совпадут с неизвестной величиной γ. Если же эти отношения заметно различны, возникает мысль о том, что в промежутке времени между сериями испытаний какие-то причины из числа $C_1, C_2, \ldots$ отпали, и возникли иные. Это изменило шансы $c_1, c_2, \ldots$, а потому и γ. Впрочем, нельзя считать подобные изменения достоверными, и мы в дальнейшем приведём выражение для их вероятности в функции $(m/\mu - m'/\mu')$ и чисел μ и μ'.

Мы вернёмся к этому следствию двух предыдущих предложений в самой теореме Якоба Бернулли, заметив, что в соответствии с гипотезой, на которой основано второе из них, дробь γ есть неизвестный шанс E, постоянный в течение обеих серий испытаний. Это событие может наступить при каждом испытании, будучи вызвано любой из причин C_1, C_2, …, действующих с одной и той же вероятностью $1/\nu$. По правилу § 5 шанс его появления, вызванного причиной C_i, равен c_i/ν, а по правилу § 10 его полный шанс равен $c_1/\nu + c_2/\nu + \ldots = \gamma$.

Для упрощения изложения мы считаем все причины равновозможными, но мы также в состоянии предположить, что каждая из них включается в их множество один или несколько раз, так что равная вероятность нарушается. Пусть каждая причина C_i включается $\nu\gamma_i$ раз, тогда γ_i выразит её вероятность и

$$\gamma = \gamma_1 c_1 + \gamma_2 c_2 + \ldots + \gamma_\nu c_\nu,$$

причём

$$\gamma_1 + \gamma_2 + \ldots + \gamma_\nu = 1,$$

потому что какая-то причина должна непременно существовать при каждом испытании. Если число причин бесконечно, вероятность каждой станет бесконечно малой. Обозначим в этом случае один из шансов c_1, c_2, …, c_ν через x, значения которого могут простираться от 0 до 1, и пусть Ydx будет вероятностью причины, которая приводит к шансу x события E. Тогда, как в § 45,

$$\gamma = \int_0^1 Yxdx, \quad \int_0^1 Ydx = 1. \qquad (52.1\text{a, b})$$

53. Пусть вместо двух возможных событий E и F их число равно λ и только одно из них наступает при каждом испытании. Так происходит, когда мы рассматриваем вещь A какой-то природы, способную принимать известные или неизвестные значения a_1, a_2, …, a_λ, из которых одно-единственное должно наступить при каждом испытании и которое является наблюдённым или будущим событием. Пусть также c_{ij} будет шансом, что причина C_i, если она достоверна, приводит к значению A, равному a_j. Числа c_{ij} с индексом i, простирающимся от 1 до ν и индексом j, от 1 до λ, известны или неизвестны, но для каждого индекса j должно выполняться равенство

$$c_{i1} + c_{i2} + \ldots + c_{i\lambda} = 1,$$

потому что при достоверности C_i одно из значений $a_1, a_2, \ldots, a_\lambda$ наверняка должно последовать.

Обозначим ещё через α_j сумму шансов a_j, которые относились или будут относиться к большому числу μ последовательных испытаний, делённую на это число; иначе говоря, α_j – средний шанс значения a_j вещи А в этих испытаниях. Полагая a_j событием Е, а множество (λ – 1) остальных значений А – противоположным событием F, мы можем принять на основе второго общего предложения § 51, что

$$\alpha_j = \gamma_1 c_{1j} + \gamma_2 c_{2j} + \ldots + \gamma_\nu c_{\nu j}.$$

Величины $\gamma_1, \gamma_2, \ldots, \gamma_\nu$ неизменно являются вероятностями различных причин, которые могут привести к появлению событий в серии испытаний, или иначе, привести к значениям вещи А, уже наблюдённым или наблюдаемым в будущем.

[3] И третье основное предложение, которое ещё оставалось за нами, состоит в том, что сумма этих λ значений А, делённая на их число, или среднее значение А, весьма вероятно лишь очень мало отличается от суммы всех своих возможных значений, умноженных соответственно на их средние шансы[16]. Обозначив через s сумму фактических значений А, с высокой вероятностью почти точно окажется, что

$$s/\mu = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_\lambda\alpha_\lambda.$$

Таким образом, если δ – сколь угодно малая дробь, всегда можно будет предположить, что μ достаточно велико, чтобы вероятность разности частей указанного равенства меньшей δ оказалась сколь угодно близкой к единице. Заметим ещё, что ввиду предыдущего выражения для α_j и соответствующих значений $\alpha_1, \alpha_2, \ldots$, правая часть не зависит от μ. Если это число очень велико, сумма s будет ощутительно пропорциональна ему. Поэтому, если обозначить через s' сумму значений А в другой серии очень большого числа μ' испытаний, разность отношений s/μ и s'/μ' весьма вероятно окажется очень небольшой. Пренебрегая ей, мы получаем

$$s/\mu = s'/\mu'.$$

В большинстве случаев число λ возможных значений А бесконечно; при возрастании от одних значений к последующим они изменяются бесконечно мало и все они заключены в заданные границы l_1 и l_2, а вероятность, которую некоторая причина C_i придаёт каждому из этих значений, поэтому бесконечно мала. Пусть $Z_i dz$ – шанс, придаваемый причиной C_i какому-то значению z. Тогда

$$\int_{l_1}^{l_2} Z_i dz = 1. \qquad (53.1)$$

Полный шанс этого значения z, или почти его средний шанс в серии испытаний, равен Zdz; для сокращения письма здесь принято, что

$$\gamma_1 Z_1 + \gamma_2 Z_2 + \ldots + \gamma_v Z_v = Z.$$

Таким образом,

$$\frac{s}{\mu} = \int_{l_1}^{l_2} Zzdz.$$

Z есть известная или неизвестная функция z. Как и каждый из интегралов (53.1), сумма дробей γ_1, γ_2, … равна единице [см. выше], и потому

$$\int_{l_1}^{l_2} Zdz = 1,$$

будь число возможных причин v конечным или бесконечным.

54. Закон больших чисел состоит в двух уравнениях

$$m/\mu = m'/\mu',\ s/\mu = s'/\mu',$$

применимых ко всем случайным физическим и моральным вещам. Его смысл следует понимать различно, в соответствии с одним или другим уравнением, и они постоянно подтверждаются, как можно заметить, разнообразными примерами, приведенными в Предисловии. Эти примеры всякого вида не могут оставить ни малейшего сомнения в общности и точности закона. Но уместно, ввиду его значимости, как необходимого основания более интересующих нас приложений исчисления вероятностей, чтобы он был доказан априорно.

Впрочем, его доказательство, основанное на предложениях §§ 52 и 53, полезно тем, что указывает нам саму причину его существования. Ввиду первого только что приведенного уравнения число m появлений события E какой-либо природы при очень большом количестве μ испытаний можно считать пропорциональным μ. В любом подобном случае отношение m/μ принимает специальное значение γ, притом точно, если μ может оказаться бесконечным, а теория указывает нам, что это значение равно сумме возможных шансов E при каждом испытании, умноженных соответственно на вероятности действовавших причин. Множество этих причин характерно отношением, существующим для каждой из них, между её вероятностью и шансом, который, будучи достоверной, она придаёт появлению E.

Пока этот закон вероятностей не изменяется, мы наблюдаем неизменность отношения m/μ в различных сериях, состоящих из большого числа испытаний. Если же, напротив, от одной серии к другой этот закон изменился, что должно было вызвать заметное изменение среднего шанса γ, мы узнаем об этом по аналогичному изменению значения m/μ. Если между этими двумя сериями наблюдений физические или моральные причины, придающие появлению E наибольшие шансы, ввиду некоторых обстоятельств стали более вероятны, значение γ возрастёт, и отношение m/μ во второй серии окажется большим, чем в первой. В противном случае это отношение уменьшится.

Если по природе события E все его возможные причины равновероятны, $Y = 1$ и $\gamma = 1/2$, так что весьма вероятно, что в длинной серии испытаний количество появлений этого события будет примерно равно половине их числа. Аналогично, если вероятности причин E пропорциональны шансам, которые они придают появлению этого события, и их число всё ещё бесконечно, то $Y = ax$, а для соблюдения условия (52.1b) должно выполняться равенство $a = 2$, так что $\gamma = 2/3$. Следовательно, при длинной серии испытаний существует вероятность, весьма близкая достоверности, что E произойдёт почти вдвое чаще, чем противоположное событие.

Однако, в большинстве случаев закон вероятностей причин нам неизвестен, средний шанс γ не может быть вычислен заранее, и лишь опыт указывает его примерное и весьма вероятное значение.

При продолжении испытаний так долго, что изменения отношения m/μ станут неощутимыми, значение γ принимается равным ему. Имея в виду все вариации шансов в течение длинной серии испытаний, почти точное постоянство отношения m/μ для событий любой природы – факт, весьма заслуживающий внимания. Можно попытаться приписать это вмешательству скрытой силы, отличной от физических или моральных причин событий, и действующей в каком-то смысле в целях порядка и неизменности. Теория, однако, доказывает нам, что, до тех пор, пока закон вероятностей причин, соответствующий каждому виду событий, не начнёт изменяться, это постоянство остаётся необходимым. В каждом случае мы должны относиться к этому как к естественному состоянию вещей, которое существует само по себе без содействия каких-либо посторонних причин; напротив, без подобной причины заметное изменение не выявляется. Это можно сравнить с состоянием покоя тел, которое существует только ввиду инерции материи, пока [и поскольку] какая-либо посторонняя причина не нарушит его.

55. Перед тем, как обратиться ко второму из предыдущих уравнений, уместно привести несколько примеров, относящихся к первому из них и подходящих для прояснения вопроса.

Пусть дано ν урн $C_1, C_2, \ldots, C_\nu$ с белыми и чёрными шариками; обозначим через c_n шанс извлечения белого шарика из какой-то урны C_n, причём тот же шанс может относиться ко многим из этих урн. Случайно выбираем одну из урн и замещаем её подобной; повторяем эту операцию несколько раз, причём множество урн неизменно остаётся без изменения. Так образуется неограниченное множество урн $B_1, B_2, \ldots$, которое состоит только из исходных урн, повторенных большее или меньшее число раз. Шанс извлечь белый шарик из B_1 обозначим через b_1, из B_2 – через b_2 и т. д. Неограниченная последовательность $b_1, b_2, \ldots$ состоит только из данных и возможно повторенных шансов $c_1, c_2, \ldots$

Из урн $B_1, B_2, \ldots, B_\mu$ извлекается по одному шарику; обозначим через β средний шанс появления белого шарика в этих μ последовательных тиражах

$$\beta = (1/\mu)(b_1 + b_2 + \ldots + b_\mu).$$

Исходные урны представляют собой ν единственно возможных причин извлечения белого шарика при каждом испытании, и если

μ – очень большое число, а m – число вышедших белых шариков, то, принимая, как и раньше, что

$$\gamma = (1/v)(c_1 + c_2 + \ldots + c_v),$$

в соответствии с предыдущим мы получаем с высокой вероятностью и почти точно, что

$$m/\mu = \beta, \beta = \gamma, m = \mu\gamma.$$

Итак, m не изменяется ощутимо при повторных тиражах из тех же урн $B_1, B_2, \ldots, B_\mu$ или из μ других урн. Если же выбрать другое очень большое число μ' урн, число извлечённых из них белых шариков весьма вероятно окажется близким $\mu' m/\mu$. Если случайно извлекать шарик μ раз подряд с возвратом из урн $C_1, C_2, \ldots$, шанс появления белого шарика окажется при всех этих испытаниях постоянным и по правилу § 10 равным γ. Если их число очень велико, число вышедших белых шариков, как и в предыдущем случае, по правилу § 49, весьма вероятно будет почти точно равно $\mu\gamma$. Эти два случая, однако, существенно различны, и их результаты совпадают только, если μ очень велико. Иначе шанс извлечь заданное число m белых шариков в первом случае будет зависеть от систем урн $C_1, C_2, \ldots$ и $B_1, B_2, \ldots$, образованных случайным образом.

Пусть исходных урн было три. Я приму $\mu = 2$ и $m = 1$ и буду определять шанс извлечения белого шарика из одной из урн B_1 и B_2 и чёрного шарика – из другой урны. Возможны девять комбинаций:

$$B_1 = B_2 = C_1; \quad B_1 = B_2 = C_2; \quad B_1 = B_2 = C_3; \quad B_1 = C_1 \text{ и } B_2 = C_2;$$
$$B_1 = C_1 \text{ и } B_2 = C_3; \quad B_1 = C_2 \text{ и } B_2 = C_3; \quad B_1 = C_2 \text{ и } B_2 = C_1;$$
$$B_1 = C_3 \text{ и } B_2 = C_1; \quad B_1 = C_3 \text{ и } B_2 = C_2$$

Для каждой из них искомый шанс имеет определённое значение, а именно, для первых трёх

$$2c_1(1 - c_1); 2c_2(1 - c_2); 2c_3(1 - c_3).$$

Для следующих трёх, и также для остальных трёх

$$c_1(1 - c_2) + c_2(1 - c_1); c_1(1 - c_3) + c_3(1 - c_1);$$
$$c_2(1 - c_3) + c_3(1 - c_2).$$

Легко заметить, что среднее значение этих девяти шансов должно равняться шансу появления белого и чёрного шариков при случайном первом извлечении из C_1, C_2, C_3 и вторичном извлечении шарика после его возвращения в ту же урну. Этот шанс равен удвоенному произведению

$$1/3(c_1 + c_2 + c_3)[1 - 1/3(c_1 + c_2 + c_3)],$$

которое в свою очередь равно 1/9 суммы предшествовавших девяти шансов.

До образования системы урн B_1, B_2, … у нас не было никаких оснований полагать, что урна B_n скорее окажется одной из урн C_1, C_2, …, а не другой. Для нас вероятность извлечь белый шарик при n-м тираже была бы поэтому равна сумме шансов $c_1 + c_2 + \ldots$, разделённой на их число, т. е. равна γ. Но, хотя она одна и та же для всех тиражей, а их число μ было сколь угодно велико, мы не посмеем заключить, в соответствии с одним только правилом § 49, что число m появлений белого шарика из урн B_1, B_2, … весьма вероятно будет очень мало отличаться от $\mu\gamma$. Мы ведь не должны упускать из вида, что это правило основано на шансе, свойственном рассматриваемому событию, а не на его вероятности или на основании полагать, что оно наступит.

56. В качестве второго примера я предположу, что очень много раз были подброшены пятифранковые монеты A_1, A_2, … Я обозначу шанс появления орла у монеты A_i, зависящий от её физического состава, через a_i. Он определяется по очень большому числу m бросков. Если он остаётся постоянным, и орёл появился n_i раз, мы можем принять по правилу § 49 $a_i = n_i/m$ за его весьма вероятное приближённое значение. Оно послужит для вычисления вероятностей различных будущих событий о бросках той же монеты. Можно держать справедливое пари, ставя m против $(m - n_i)$, на то, что орёл появится при следующем броске; m^2 против $(m^2 - n_i^2)$, что он появится дважды подряд; $2n_i(m - n_i)$ против $m^2 - 2n_i(m - n_i)$, – что только один раз в двух бросках и т. д.

В новой серии очень большого числа m' испытаний число n_i' появлений орла с высокой вероятностью, опять же по правилу § 49, окажется почти равным $m'a_i$. Отношения n_i/m и n_i'/m' должны очень мало отличаться друг от друга. Но всё-таки, поскольку это значение a_i получено опытным путём, оно лишь весьма вероятно, но не достоверно. Вероятность малой разности между этими отношениями не так высока, что будет видно в дальнейшем, как когда этот шанс a_i достоверен и задан заранее.

Пусть вместо многократных бросков одной и той же монеты последовательно подбрасывается очень большое их число μ, случайно выбранных из всех монет одного и того же способа чеканки. Пусть, далее, орёл появился n раз, а средний шанс этого события, не только для подброшенных монет, но для всех монет того же вида и того же способа чеканки, равен α. Ввиду двух общих предложений § 52 весьма вероятно, что почти точно $\alpha = n/\mu$, как если бы все неизвестные шансы $a_1, a_2, \ldots$ были равны друг другу.

При $n/\mu > 1/2$ и $< 1/2$ мы заключаем, что, соответственно, для пятифранковых монет этой чеканки шанс появления орла, вообще говоря, больше или меньше шанса другого возможного исхода. Для монеты A_i шанс a_i отличается от α, и возможно, что $n_i/\mu < 1/2$, хотя $n/\mu > 1/2$ или наоборот. Снова подбрасывая те же монеты, или, вообще, очень большое число μ′ других подобных монет с теми же изображениями, величина α, определённая выше, не изменится. Поэтому, если орёл появится n' раз, то

$$n'/\mu' = n/\mu, \qquad (56.1)$$

а для одной и той же монеты A_i $n_i/m = n_i'/m'$. Однако, подобно изменениям n_i/m от одной монеты к другой, равенство (56.1) не будет всегда удовлетворяться, если подброшенные монеты не были одного и того же вида и одной и той же чеканки.

57. Постоянные шансы и средние шансы событий определяются по опыту одним и тем же способом и с одной и той же вероятностью[17], однако имеются существенные различия в их возможных приложениях. И те, и другие шансы позволяют сразу же устанавливать вероятность появления рассматриваемого события при новом одиночном испытании, но обстоятельства меняются, если изучается составное событие.

Если дело идёт о совпадении результатов двух последовательных бросков пятифранковой монеты, то следует рассмотреть два различных случая. Можно предположить, что были подброшены либо различные или те же самые монеты, случайно выбранные из всех λ монет $A_1, A_2, \ldots$ одной и той же чеканки, либо одна и та же монета, также выбранная случайно. В первом случае вероятность совпадения исходов зависит только от α, введенного в § 56, и то же происходит при постоянных шансах.

Во втором случае вероятность, помимо прочего, зависит от другой неизвестной величины, на которую она отличается от своего значения при неизменных шансах. Чтобы прояснить это, я замечу, что относительно двух каких-либо монет A_i и A_j вероятность совпадения результатов в двух последовательных испытаниях равна

$$a_i a_j + (1 - a_i)(1 - a_j).$$

В первом из двух случаев каждая монета A_1, A_2, … могла быть объединена или сама с собой, или с любой из остальных, и число всех этих равновозможных комбинаций равно λ^2. Обозначим полную вероятность совпадений через s; тогда, по правилу § 10,

$$s = (1/\lambda^2)[\sum a_i \sum a_j + \sum(1 - a_i)\sum(1 - a_j)],$$

причём суммы распространяются от $i, j = 1$ до $i, j = \lambda$.

Пусть

$$\alpha = 1/2(1 + k),\ a_i = 1/2(1 + k + \delta_i),\ a_j = 1/2(1 + k + \delta_j),$$

где k, δ_i, δ_j, $k + \delta_i$, $k + \delta_j$ обозначают положительные или отрицательные дроби. Первая из них выводится из отношения n/μ из § 56, которое известно по наблюдениям, остальные же изменяются от одной монеты к другой так, чтобы

$$\sum\delta_i = 0,\ \sum\delta_j = 0$$

и в то же время

$$1 - a_i = 1/2(1 - k - \delta_i),\ 1 - a_j = 1/2(1 - k - \delta_j).$$

С учётом предыдущих уравнений суммы в выражении для s оказываются равными

$$\sum a_i = \sum a_j = 1/2\lambda(1 + k),\ \sum(1 - a_i) = \sum(1 - a_j) = 1/2\lambda(1 - k)$$

и потому

$$s = 1/2(1 + k^2),$$

т. е. зависит лишь от k или среднего шанса α появления орла, но никак не от неравенств шансов δ_1, δ_2, …

Если снова подбрасывать две случайно выбранные монеты очень большое число раз a, s окажется средним шансом повторения исходов в этой серии сдвоенных испытаний. Пусть совпадения произошли b раз, тогда, по § 52, приближённо $b = as$, что можно проверить по опыту.

Во втором случае, когда каждая пара испытаний производится с одной и той же монетой A_i, вероятность повторения исходов равна

$$a_i^2 + (1 - a_i)^2$$

и полная вероятность совпадений равна

$$s' = (1/\lambda)[\sum a_i^2 + \sum(1 - a_i)^2] = 1/2(1 + k^2 + h^2),\ h^2 = (1/\lambda)\sum\delta_i^2.$$

Эта вероятность превышает вероятность s первого случая и зависит от новой неизвестной h, которая в свою очередь зависит от неравенств между δ_1, δ_2, …

Если снова очень много раз a' дважды подбрасывать одну и ту же случайно выбранную монету, то s' выразит вероятность повторения исходов в этой серии сдвоенных испытаний. И если повторения произошли b' раз, то почти точно $b' = a's'$, откуда можно определить значение h, тогда как k уже было известно.

58. Замечу, что если подбросить трижды подряд одну и ту же монету, случайно выбранную из A_1, A_2, …, вероятность совмещения исходов выразится через вероятность s' и поэтому станет известна без проведения дополнительных испытаний. Относительно некоторой монеты A_i эта вероятность равна

$$a_i^3 + (1 - a_i)^3,$$

а полная вероятность окажется равной

$$s'' = \frac{1}{\lambda}[\sum a_i^3 + \sum(1 - a_i^3)] = \frac{1}{4}[1 + 3(k^2 + h^2)] = \frac{1}{2}(3s' - 1).$$

Эта величина s'' будет также средним шансом совпадения исходов в очень длинной серии строенных испытаний. Если обозначить их число через a'', а через b''– число совпадений, то $b'' = a''s''$.

Подставляя в предыдущее уравнение b'/a' и b''/a'' вместо s' и s'', мы определим соотношение

$$a'a'' = 3b'a'' - 2b''a',$$

которое окажется тем вероятнее и тем точнее, чем больше величины a', a'', b', b''.

Она [s''?] не зависит от закона величин a_1, a_2, … и поэтому останется без изменения при равенстве указанных величин друг другу, т. е. при применении одной и той же монеты в каждом сдвоенном или строенном испытании. Таким образом, если подбросить одну и ту же монету очень много раз, именно $6c$ раз, разделить полученную серию однократных испытаний на сдвоенные (первое из них будет состоять из первых двух испытаний; второе – из третьего и четвёртого и т. д.) и строенные (первое, соответственно, из первых трёх испытаний, второе – из следующей тройки и т. д.), и применить к полученным сериям предыдущее уравнение, мы получим

$$a' = 3c,\ a'' = 2c,$$

так что $c = b' - b''$. Аналогичные соотношения можно вывести для случаев большего числа совместно рассматриваемых испытаний.

59. Эти соображения можно непосредственно применить к мужским и женским рождениям, для чего достаточно заменить монеты женитьбами и принять шанс мужского рождения в женитьбе $А_i$ за a_i.

Во Франции общее число ежегодных рождений почти достигает миллиона. Наблюдения указывают, что отношение мужских рождений к женским превышает единицу примерно на 1/15. В течение 10 лет, с 1817 по 1826 гг., его среднее значение было равно 1.0656, а его крайние значения едва уклонялись от среднего на 0.005 в ту или иную сторону. На этих наблюдениях были основаны результаты моего мемуара (1830). За 1817 – 1833 гг. включительно это среднее отношение равнялось 1.0619, что не отличалось от предыдущего более, чем на 0.005.

Причина превышения мужских рождений нам неизвестна; можно полагать, что она сильно изменяется от одной женитьбы к другой и что шансы a_1, a_2, … весьма различны, так что многие из них без сомнения опускаются ниже 1/2. Однако, как было указано, за 17 лет соотношение годичных рождений полов очень мало изменилось, что является примечательным подтверждением *закона больших чисел*. Если принять 16/31 за отношение большого числа мужских рождений к соответствующему числу всех рождений, то оно же будет средним шансом мужского рождения, так что k из § 57 принимает значение 1/31.

Мы оставляем в стороне вопрос о том, остаётся ли шанс мужского рождения тем же самым для каждого младенца, родившегося в одной и той же женитьбе, или же он изменяется, допустим, так же, как от одной женитьбы к другой. Во втором случае средний шанс повторения пола в двух первых рождениях примерно равен $1/2(1 + k^2)$ и превышает 1/2 лишь почти на полу-тысячную. Поэтому количество подобных повторений в очень большом числе двух первых детей превышает половину этого числа на полу-тысячную.

В первом случае первое из этих двух чисел может намного превысить половину второго ввиду неизвестной величины h, которая входит в выражение $1/2(1 + k^2 + h^2)$ среднего шанса совпадения. Для очень большого числа женитьб материал § 58 будет неизменно применим к количеству повторений пола при двух и трёх первых рождениях.

60. Ввиду второго уравнения § 54, если A – какая-то вещь, которая в каждом испытании может принимать различные значения, сумма её значений в очень длинной серии испытаний весьма вероятно будет почти точно пропорциональна их числу. Для данного A отношение этой суммы к указанному числу будет по мере его дальнейшего возрастания неограниченно стремиться к особому значению, зависящему от закона вероятностей различных возможных значений A, и достигнет его, если число испытаний может стать бесконечным.

По поводу этого отношения можно заметить подобное тому, что было указано в § 54 при рассмотрении первого уравнения. Второе уравнение, или скорее

$$\frac{s}{\mu} = \int_{l_1}^{l_2} Zzdz,$$

приводит, как и первое, к многочисленным полезным приложениям.

Пусть, к примеру, α – угол, который желательно измерить. Он существует, его величина однозначна и определённа. Но, ввиду неизбежных и переменных ошибок наблюдения, при каждом измерении этот угол может принять бесконечное множество значений. Я полагаю, что этот многократно и последовательно измеренный угол является вещью А, так что Zdz выразит шанс её некоторого значения z, которое следует из конструкции инструмента и умения наблюдателя.

Обозначим через k абсциссу центра тяжести площади плоской кривой, которая простирается по оси абсцисс от $z = l_1$ до l_2, т. е., как в § 53, заключена в границах возможных значений А, и через Z – его ординату. Примем

$$z = k + x,\ l_1 = k + h_1,\ l_2 = k + h_2$$

и пусть X будет значением Z при $z = k + x$. Тогда

$$\int_{l_1}^{l_2} Zdz = \int_{h_1}^{h_2} Xdx = 1,\ \int_{h_1}^{h_2} Xxdx = 0$$

и поэтому, ввиду указанного уравнения, в соответствии с которым s является суммой значений А при большом числе μ испытаний, почти точно $s/\mu = k$.

Именно к постоянной k стремится, стало быть, её среднее значение s/μ по мере дальнейшего возрастания μ. Но даже когда это отношение становится ощутимо постоянным, т. е. когда оно ощутимо неизменно, для многих других серий из большого числа измерений может иногда случиться, что это среднее будет намного отличаться от определяемого угла α; оно всегда будет приближённым значением константы γ, которое, возможно, вовсе не совпадает с этим углом[18].

Действительно, пусть

$$z = \alpha + u,\ l_1 = \alpha + g_1,\ l_2 = \alpha + g_2.$$

Обозначим через U значение Z при замене z на $(\alpha + u)$. Тогда

$$\int_{l_1}^{l_2} Zdz = \int_{g_1}^{g_2} Udu = 1,\ k = \alpha + \int_{g_1}^{g_2} Uudu.$$

Разность u между углом α и некоторым возможным значением z измеренного угла А равна возможной ошибке инструмента и наблюдателя. Она может быть положительна и отрицательна и простираться от $u = g_1$ до g_2 и её бесконечно малая вероятность равна Udu. И если ввиду устройства инструмента не возникает никакой причины, придающей положительным ошибкам больше шансов, чем отрицательным, или отрицательным – больше, чем положительным, и если то же имеет место в работе исполнителя, числа g_1 и g_2 будут равны друг другу по значению и противоположны по знаку, а функция U будет принимать равные значения при равных и противоположных по знаку значениях переменной u. Тогда

$$\int_{g_1}^{g_2} U u du = 0,\ k = \alpha.$$

В этом наиболее частом случае отношение s/μ окажется приближённым значением α. Но если либо устройство инструмента, либо метод визирования, который применяет наблюдатель, придаёт некоторый перевес положительным или отрицательным ошибкам, предшествующий интеграл не будет нулём, константы α и k не совпадут и отношение s/μ, вообще говоря, заметно отклонится от истинного значения α. Заметить это обстоятельство можно только, если измерить тот же угол либо другим инструментом, либо поменяв наблюдателя. Я ограничиваюсь указанием на это приложение исчисления вероятностей. По поводу ошибок наблюдения и методов вычисления, подходящих для уменьшения и оценивания их влияния, я отсылаю к Лапласу (Laplace 1812) и своим мемуарам (1824; 1829).

61. В качестве второго примера на уравнение, приведенное в начале § 60, я предположу, что причины C_1, C_2, … определяют шансы продолжительности человеческой жизни в определённую эпоху в какой-либо стране. Помимо прочих, этими причинами являются физическая конституция новорожденных, благополучие населения, болезни, ограничивающие эту продолжительность, а также несомненно те, которые возникают в самой жизни и препятствуют её продолжению за никогда не пересекавшиеся пределы. Можно полагать, что, будь болезни единственной причиной смерти, и что они, так сказать, случайны, некоторые из громадного числа живших избегали бы эти опасности в течение более двух столетий [и никогда бы не умерли], но подобные факты никогда не были отмечены.

Вещью А окажется здесь продолжительность жизни новорождённого, возможным значением А – z, а Zdz будет этим шансом z , возникающим ввиду действия всех причин, каковы бы они ни были, которые могут определить z, но не для определённого младенца, а для рода человеческого в данных эпоху и месте. Итак, представим, что определённая физическая конституция при рождении придаёт шанс $Z'dz$ прожить в точности z, что другая конституция придаёт шанс $Z''dz$ прожить столько же и т. д. Пусть ς', ς'', … будут вероятностями этих различных конституций; тогда, ввиду указанных причин, $Z = Z'\varsigma' + Z''\varsigma'' + \ldots$ Сумма распространяется на все возможные конституции, а если их число бесконечно, Z станет определённым интегралом с неизвестным, но установленным значением. В стране, в которой новорожденные крепче или обладают лучшей конституцией, этот интеграл несомненно будет бо́льшим. В данной стране он может не быть одним и тем же для обоих полов. Кроме того, несомненно, что значения Z', Z'', … зависят от возможных болезней и благополучия населения. Функция Z переменна, а потому и интеграл от $Zzdz$ тоже изменит своё значение между двумя удалёнными друг от друга эпохами, если за это время исчезнет какое-то заболевание или если с прогрессом общества улучшится благосостояние жителей.

Можно, если угодно, принять в этом интеграле пределы $l_1 = 0$ и $l_2 = \infty$ и считать Z функцией, исчезающей по ту сторону от какого-то z, которое не известно так же, как и форма Z. И наблюдёнными значениями А будут возрасты, в которых умерло очень большое число μ людей, рождённых в той же стране и в ту же эпоху. Обозначим сумму этих возрастов через s, тогда весьма вероятно, что почти точно

$$\frac{s}{\mu} = \int_0^\infty Zzdz.$$

Следовательно, это отношение s/μ или то, что называется *средним сроком жизни*, остаётся постоянным для данной страны, пока какая-либо из известных или неизвестных причин C_1, C_2, … не испытает заметного изменения.

Во Франции предполагается, что средний срок жизни составляет 29 лет, однако эта оценка основана на наблюдениях, проведенных до применения вакцины [против оспы], и уже сильно устарела. Сегодня он должен быть намного длиннее, и следовало бы пожелать его определения заново, отдельно для мужчин и женщин, для различных состояний и разнообразных территорий королевства.

Средний срок жизни изучается также для данного возраста, и тогда s – число остающихся лет жизни для очень большого числа μ лиц. Отношение s/μ – средний срок жизни для этого возраста; он меняется вместе с возрастом, но постоянен для данного возраста. Предполагается, что он достигает наибольшего значения при возрасте между четырьмя и пятью годами, доходя для него до 43 лет. Другим средством изучения срока жизни являются таблицы *смертности*, составленные для очень большого числа μ лиц, рождённых в одной и той же стране и в ту же самую эпоху. Они указывают число доживающих после года, двух, трёх, … лет вплоть до полного вымирания. Пусть число доживших до данного возраста равно m. Ввиду первого уравнения § 54 отношение m/μ ощутимо постоянно, по крайней мере до преклонного возраста, когда m станет весьма малым. К ста годам, например, это постоянство заключается в том, что отношение m/μ неизменно оказывается очень малой дробью.

В интеграле

$$\int_0^\infty Zzdz$$

вместо бесконечно малых приращений z можно полагать, что он возрастает на весьма малые интервалы. Если для определённости каждый из этих интервалов времени считать единичным по времени и обозначить через h_1, h_2, … ряд значений z, а через H_1, H_2, … – соответствующие значения Z, то, как известно, сумма произведений H_1h_1, H_2h_2, … будет приближённым значением этого интеграла. Средний срок жизни новорожденного окажется равным

$$v = H_1h_1 + H_2h_2 + \ldots$$

Здесь H_n – шанс умереть в возрасте h_n, и таким образом средний срок жизни v можно считать математическим ожиданием (§ 23) [этого срока для] новорожденного с неизвестной физической конституцией. Впрочем, в соответствии с таблицами смертности, составленными для очень большого числа младенцев, более половины из них умирает до достижения возраста v.

62. В качестве последнего примера предположим, что в данной местности и в данный день года вычислен избыток высоты уровня моря, возникший под одновременным действием Солнца и

Луны. За вещь А примем разность между этим избытком и наблюдённым в том же месте и в ту же эпоху каждого года. Значения А изменяются из года в год ввиду ветров, которые могут веять в том месте в эту эпоху и которые определяют шансы этих различных значений.

Если рассматривать все возможные направления и силы этих ветров, их соответствующие вероятности и шансы, придаваемые этими причинами значениям z вещи А, то интеграл

$$\int_{l_1}^{l_2} Zz dz$$

примет неизвестное, но определённое значение, которое останется неизменным, пока закон вероятностей каждого возможного ветра не изменится. Отношение s/μ тоже будет почти неизменным; s здесь является суммой значений А, наблюдённых в течение многих лет.

Мы заранее не знаем, равно ли s/μ нулю или ничтожной дроби, т. е. не знаем, пренебрежимо ли влияние ветров на общие законы *приливов*. Лишь опыт может указать значение этого отношения и сообщить нам, изменяется ли оно в различные эпохи года и от одного места к другому, в которых (на морских побережьях, в портах и в отрытом море) были проведены наблюдения. Для установления влияния определённого ветра следует использовать только значения А, наблюдённые при его действии. Чтобы не нуждаться в очень большом числе лет наблюдения, эти значения можно всё-таки получить из многих последовательных дней, в течение которых направление ветра мало изменялось. Многие учёные ныне занимаются этими исследованиями, которые требуют длительного труда и приводят к интересным результатам.

63. Изложение правил исчисления вероятностей и их общих следствий в этой и предыдущей главах закончено, и я возвращаюсь к понятиям *причин* и *следствий*, которые были лишь намечены в § 27.

Причина, свойственная вещи Е, – это, как было сказано в § 27, другая вещь С, обладающая *силой*, наверняка приводящей к Е, каковы бы ни были её природа и метод её действия. Так, то, что называется притяжением Земли, является определённой вещью, сила которой заставляет падать не поддерживаемые тела, находящиеся на её поверхности. Аналогично, по нашему желанию и при помощи наших мышц и нервов мы можем произвести часть тех движений, которые поэтому называются волевыми.

Иногда в природе существует лишь одна причина С, которая способна произвести вещь А, так что наблюдение Е всегда предполагает вмешательство С. В других случаях эта вещь может быть приписана многим отличающимся друг от друга причинам, которые действуют совместно или исключают друг друга, и тогда только одна из них приводит к Е. Таковы, как я полагаю, простейшие понятия о принципе причинности. Тем не менее, знаменитый историк Англии выразил иное мнение об этой метафизической теме, которая заслуживает более подробного исследования и на которую исчисление вероятностей может пролить немало света.

По Юму (1748), понятие *причинности* мы можем понимать лишь как *сочетание*, но не как необходимую *связь* между тем, что мы называем *причиной* и *следствием*. И это сочетание является для нас только сильным предположением, результатом многократно замеченного. Наблюдая что-то небольшое число раз и предполагая, что и впредь произойдёт то же, мы, стало быть, судим о природе по слишком незначительному образцу. Другие придерживаются того же мнения и пытаются обосновать его правилами о вероятности будущих событий по существующим наблюдениям[19].

Но Юм идёт дальше. Даже без обращения к этим законам вероятности он полагает, что привычка видеть следствие вслед за причиной производит в нашем сознании некоторую ассоциацию идей, и мы начинаем верить, что следствие имеет место тогда, когда есть причина. Так действительно может происходить у большинства людей, не исследующих принцип своей веры и её степень вероятности. Для них, это соединение идей следует сравнивать с тем, что в нашем сознании, вне зависимости от наших мыслей и желания, название вещи напоминает саму вещь.

Одним из примеров, которые автор избрал для изложения своего мнения, был удар движущегося тела о свободное тело, находящееся в покое, и движение второго тела после столкновения с первым. Мы многократно замечаем подобное сочетание удара и движения тела, подвергнувшегося ему, но ни разу не наблюдали противоположного события. Этого достаточно, чтобы мы, отвлекаясь от любых иных соображений, получили серьёзное основание верить, что очень высокая вероятность указанного сочетания означает, что оно будет происходить и впредь. То же самое происходит по поводу всех сочетаний причин и следствий, которые мы без исключения ежедневно наблюдаем. Их вероятности подпитываются, так сказать, этим непрерывным опытом, и в соответствии с привычкой разум или вычисление сильно уверяют нас, что в будущем за этими причинами всегда произойдут их следствия.

Но если мы недостаточно часто наблюдали некоторое явление вслед за приписываемой ему причиной, то, в соответствии с ранее изложенными правилами, будущее сочетание этой причины и этого следствия обладает не слишком высокой вероятностью. Тем не менее, мы редко сомневаемся в появлении этого явления вновь, если снова имеет место его причина. Подобная уверенность предполагает, что наш разум приписывает причине некоторую *силу* производить своё следствие и допускает необходимую связь между указанными понятиями вне зависимости от наблюдённого числа их сочетаний, будь оно велико или нет.

Так, Эрстед обнаружил, что при соединении полюсов вольтова столба металлической проволокой свободно подвешенная неподалеку магнитная стрелка отклонялась от своего естественного направления. После немногочисленного повторения этого фундаментального опыта знаменитый физик без сомнения уверился, что указанное явление будет несомненно воспроизводиться и впредь. Тем не менее, если наше основание верить в подобное воспроизведение зиждется лишь на сочетании вольтова столба и отклонения магнитной стрелки, наблюдённого, скажем, десяток раз, вероятность, что это явление повторится при новом испытании, окажется равной лишь 11/12 (§ 46). Можно поставить 11 против 10 за то, что в 10 опытах это же явление повторится без перерыва, но разумно полагать, что в более длинной последующей серии опытов оно не будет повторяться всегда.

Я дополнительно приведу для примера недавнее удачное применение указанного к химическому составу тел, которое в определённом смысле произвёл Био, наблюдая *последовательную поляризацию света*, существование которой в однородных и не кристаллических средах было давно ему известно[20]. После немногих тщательных наблюдений устанавливается, что данное вещество отклоняет поляризованный луч, скажем, вправо от наблюдателя и что отмеченные отклонения были достаточно велики и не оставили никакой неопределённости в их направлении.

Этого достаточно, чтобы уверить нас так же, как в существовании какой-то вещи, в которой никто не сомневается, что и впредь то же вещество будет всегда отклонять свет вправо. Тем не менее, сочетание этого вещества и отклонения вправо, отмеченное при немногочисленных наблюдениях, придаёт лишь слабую вероятность, даже меньшую 1/2, что при подобном или большем числе новых испытаний произойдёт отклонение вправо. Эти и аналогичные воображаемые примеры доказывают, как мне кажется, что доверие нашего разума в повторении следствий после своих причин не может основываться лишь на прежних более или менее многочисленных наблюдениях. Заметно, что, вне зависимости от какой-либо привычки нашего разума, одна-единственная возможность определённой способности причины наверняка производить следствие намного повышает основание веры в подобное повторение и может подвести его вероятность очень близко к достоверности, даже если предыдущие наблюдения были немногочисленными.

64. Перед наблюдением явления P, если известно, что оно появится или нет в целой серии предстоящих испытаний, мы допускаем, что существование причины C, способной наверняка привести к нему, не является невозможным. Ввиду определённых обстоятельств мы также заранее понимаем, что её существование обладает некоторой вероятностью p, и что она становится более или менее правдоподобной. Предположим также, что P наблюдалось во всех n испытаниях, после чего вероятность существования C изменилась и стала равной w; требуется определить её.

Как бы тщательно мы не старались уменьшить влияние других причин, способных при отсутствии C привести к явлению P при каждом испытании, можно полагать, что это влияние всё же не исчезло. Пусть, стало быть, существуют определённые причины B_1, B_2, …, B_n, известные или нет, которые при отсутствии C также могут привести к этому явлению при сочетании со случайностью (§ 27), а именно, B_1 приводит к нему в первом испытании, B_2 – во втором, …, B_n – в последнем. Пусть вообще r_i будет вероятностью существования B_i, умноженной на шанс того, что она, если достоверна, приведёт к явлению P, и примем для сокращения письма

$$r_1 r_2 \dots r_n = \rho.$$

Это произведение является вероятностью наступления Р во всех n испытаниях ввиду действия множества причин B_1, B_2, … при отсутствии С. Поскольку $1 - p$ есть вероятность того, что С не существует, то, если принять эту гипотезу, $(1 - p)\rho$ окажется вероятностью наблюдённого события, т. е. постоянного появления Р. При противном предположении его вероятность равна p, т. е. вероятности существования С до наблюдений, потому что эта причина наверняка приводит к Р во всех испытаниях. Следовательно, по правилу § 28 вероятность этой второй гипотезы или существования С после наблюдения равна

$$w = \frac{p}{p + (1-p)\rho},$$

а вероятность её отсутствия

$$1 - w = \frac{(1-p)\rho}{p + (1-p)\rho}.$$

К этому же результату можно придти, принимая во внимание n испытаний последовательно, а не зараз. В соответствии с гипотезой вероятность существования С до первого испытания была p; после первого испытания, p'; после второго – p'' и т. д. Тогда

$$p' = \frac{p}{p + (1-p)r_1},\ 1 - p' = \frac{(1-p)r_1}{p + (1-p)r_1},$$

$$p'' = \frac{p'}{p' + (1-p')r_2},\ 1 - p'' = \frac{(1-p')r_2}{p' + (1-p')r_2},\ldots$$

Исключая p' и $(1 - p')$ из значений p'' и $(1 - p'')$, затем исключая p'' и $(1 - p'')$ из значений p''' и $(1 - p''')$ и т. д., мы получим предыдущие выражения w и $(1 - w)$ для вероятностей существования и отсутствия С после n испытаний. Далее, пусть w' будет вероятностью постоянного наступления явления Р в новой серии из n' испытаний. Каково бы ни было это число n', вероятность указанному результату произойти под действием причины С, если она достоверна, равна вероятности w существования С, выведенной из первых n испытаний. При отсутствии этой причины наступление Р могло также произойти ввиду других причин $B'_1, B'_2, \ldots, B'_{n'}$, подобных $B_1, B_2, \ldots, B_n$, чьё влияние нельзя было полностью исключить. Соответственно, предыдущие величины $r_1, r_2, \ldots, r_n$ станут равными $r'_1, r'_2, \ldots, r'_{n'}$, так что r'_i по отношению к B'_i является тем же, что и r_i по отношению к B_i. Пусть также

$$r'_1 r'_2 \ldots r'_{n'} = \rho'.$$

Если причина С не существует, то вероятность появления Р в новых испытаниях числом n' окажется равной $(1 - w)\rho'$ и мы заключаем, что полным выражением для w' будет

$$w' = w + (1 - w)\rho'.$$

Подставляя вместо w и $(1 - w)$ их предыдущие значения, мы получим

$$w' = \frac{p + (1-p)\rho\rho'}{p + (1-p)\rho}.$$

Выражения для w и w' указывают, что вероятность существования С, которая возможно была очень низкой до наблюдения Р, может стать очень высокой после немногочисленных наблюдений этого явления и придать вероятность, весьма близкую к достоверности, его постоянному появлению в следующих испытаниях. Пусть, например, по каким-то причинам, и, если угодно, ввиду предубеждения, априорная вероятность С равнялась лишь 1/100 000. Примем также, что влияние случайных причин, несмотря на меры, принятые для их исключения, могут ещё быть такими, что величины $r_1, r_2, \ldots$ равняются 1/10 или меньшими, и, наконец, что Р наблюдалось без перерыва лишь 10 раз. Тогда

$$p = 0.00001,\ \rho < 0.00001p,\ w > \frac{1}{1 + 0.00001(1-p)}.$$

Поэтому после наблюдений вероятность существования С будет отличаться от единицы менее, чем на 1/100 000, и её отсутствие станет ещё менее вероятным, чем её существование до наблюдений. Каково бы ни было число n', вероятность w' того, что Р будет постоянно наступать в новой серии n' испытаний, станет ещё выше, чем вероятность существования С, или не сможет оказаться ниже.

65. В описанном приложении исчисления вероятностей причина С рассматривалась отвлечённо, т. е. вне зависимости от какой-либо теории, подводящей явление Р к более общим законам и обеспечивающей точное изложение в соответствии с причиной, приписываемой Р, что ещё более повысит её вероятность. Мы рассматривали это явление Р как происходившее постоянно, и предшествующие вычисления имели целью указать, как наша вера в его предстоящее появление по небольшому числу его наблюдений не могла быть основана лишь на идее существования причины, способной наверняка приводить к подобным явлениям. Впрочем, исчисление вероятностей не может ни указать нам, какова эта действенная причина, ни определить, какая причина наиболее вероятна среди многих, способных наверняка привести к Р и потому возможно приписанных этому явлению.

Если явление Р не появляется в одном или многих испытаниях, и если тем не менее существует уверенность, что причина С, если она существует, способна наверняка привести к нему, и что она должна была бы действовать во всех этих испытаниях, то очевидно, что ни она, ни какая-либо другая причина того же вида не существует. Но помимо подобных причин имеются иные, которые действуют во всех испытаниях, и они лишь способны придать определённый шанс появлению некоторого явления Р (§ 27) при сочетании со случайностью, или с переменными причинами, которые действуют лишь иногда.

Эти переменные и иррегулярные причины, которые мы не должны смешивать со случайностью, могут повлиять на средний шанс наступления Р в длинной серии испытаний и поэтому (§ 52) на число их появлений или предстоящих появлений, делённое на общее количество испытаний. Но если принять меры, чтобы по возможности уменьшить влияние случайных причин, и можно было бы считать их по существу исчезнувшими, и если явление Р наступало m раз в очень большом числе μ испытаний, вероятность того, что существует постоянная причина, либо благоприятная для этого результата, либо нет в зависимости от того, будет ли m заметно больше или меньше половины μ, окажется очень высокой.

Примем для примера случай тела с двумя гранями, подброшенного очень много раз. Если исходы бросков заметно отличаются друг от друга, как в эксперименте Бюффона (§ 50)[21], существование причины, благоприятной или нет появлению определённой грани, можно считать исключительно вероятным. Какова эта постоянная причина? Исчисление вероятностей лишь сообщает, что она необходима, но не способно указать её природу. Именно законы механики дают нам знать, каково неравенство весов в некоторой части подбрасываемого тела, но, ввиду сложности этой проблемы, не определяют ни влияния подобной причины, ни шанса, которого она придаёт появлению каждой грани, что можно выяснить только по опыту.

Подобным образом можно определить, как предложил Лаплас (Laplace 1814, с. 133), существование или отсутствие каких-либо скрытых причин, которые заранее не были совершенно невозможны и которые не в состоянии наверняка приводить к соответствующим явлениям. Для достижения этой цели требуются длинные серии испытаний, чтобы по возможности исключить влияние случайных причин, и кроме того необходимо точно подсчитать количества появлений и отсутствия изучаемого явления. Если отношение первого их этих чисел ко второму заметно превысит единицу, существование некоторой причины и шанса, который она придаёт этому явлению, становятся весьма вероятными.

Если игроки А и В сыграли очень большое число μ партий, из которых А выиграл m и если m/μ превысит 1/2, притом не на очень небольшую дробь, существование причины, благоприятной А, можно считать почти достоверной. Если ни один игрок не даёт фору другому, этой причиной следует считать превосходство А над В, а его, так сказать, мерой является m/μ.

При игре в пикет, к примеру, результат каждой партии зависит только от разности умений игроков и распределения карт между ними. Если никто из них не жульничает, это распределение определяется случаем и может повлиять на соотношение количества выигранных каждым игроком партий, если, однако, этих партий было немного. Это то, что можно назвать *удачей* и *неудачей*, если только не приписывать этих понятий самим игрокам. Действительно, нелепо предполагать, что существует какое-то соотношение между ними и картами, которые распределяются чисто случайно. Каждый раз те, которые выпадают одному игроку, могли бы так же само достаться другому. Но в достаточно длинной серии партий на шансы игры может повлиять только различие в умении игроков.

В перспективе счастливы только умелые игроки, и наоборот. Если А и В вновь играют очень большое число μ' партий, с очень высокой вероятностью количество выигрышей А окажется почти равным $\mu' m/\mu$. В противном случае можно было бы обоснованно полагать, что в интервале между сериями игры превосходство А над В увеличилось или уменьшилось.

Примечания

1. См. Прим. 4 к гл. 1.

2. Таким образом, событие Е, как и все иные, вводимые автором, являются случайными.

3. Но обязательно ли *множество* причин? Это определение можно было бы понимать как случайность в широком смысле, не подчиняющуюся никаким вероятностным законам и потому не изучаемую теорией вероятностей (Колмогоров 1983). Примерами могут служить уклонения в схеме лестницы существ (Lamarck 1815, с. 133). Однако, оказывается, что случайность действует только на случайные события, так что определение Пуассона явно неудачно.

4. Fienberg (1971) заметил, что, даже вне зависимости от мнения Пуассона, трудно добиться равенства шансов извлечения шариков из урны.

5. См. Прим. 18.

6. Ввиду допущенной по этому поводу ошибки решение задачи в § 17 моего мемуара (1830) неточно, и я вывел из него неверное следствие. Автор.

7. См. Прим. 3 к Предисловию.

8. Подобный же пример встречается у Лапласа, см. Прим. 9 к гл. 1.

9. Известно, что в конце XIX века физики, например, вообще считали, что в скором времени в их науке не останется ничего нового. Жизнь опровергла подобные мнения.

10. Ср. Гаусс (1823, §§ 18 и 19): независимые линейные функции наблюдений не должны содержать общих наблюдений.

11. Автор несколько раз (также и в гл. 4) использует эту неточную фразу.

12. Конец этой фразы мы можем только привести в оригинале и в его немецком переводе издания 1841 г.:

la probabilité de similitude dans deux épreuves consécutives, est la même que s'il y avait, entre les chances de G et de l'événement contraire, une différence $1/\sqrt{3}$, *sans que l'on connût la chance la plus favorable*.

so ist die Wahrscheinlichkeit der Übereinstimmung zwei auf einander folgenden Versuche dieselbe, als wenn zwischen der Wahrscheinlichkeit der Ereignisse G und der des entgegengesetzten Ereignisses eine Differenz $1/\sqrt{3}$ *stattfinde, ohne daß man der günstigste Wert kennt.*

13. В Прим. 10 к гл. 1 мы уже заметили, что автор приводил численные данные с явно избыточными значащими цифрами.

14. Следовало бы пояснить, как λ стало вероятностью.

15. Неясно, какие именно уравнения имел в виду автор. Вообще же последняя фраза непонятна, а чуть ниже он указал, что теорема Якоба Бернулли основана на интегральном исчислении, хотя Бернулли обошёлся элементарными средствами.

16. Автор часто применял математическое ожидание не ссылаясь на него.

17. Неясно, как те и другие шансы определяются с *одной и той же* вероятностью.

18. При описании измерения угла автор допустил неточности и ошибку. *Определённое*, “истинное” значение угла можно понимать только в смысле Фурье (и многих последующих авторов) как предел среднего арифметического при бесконечном числе измерений (Шейнин 2007). Ошибки наблюдения происходят также от метеорологических условий. И, вопреки автору, наблюдения практически всегда искажаются систематическими ошибками.

19. Вот мнение Юма (1739/1969, Bk. 1, pt. 3, § 11, p. 124): *Нелепо будет выглядеть тот, кто скажет, что завтрашний восход Солнца лишь вероятен*. Задача о вероятности этого события стала классической.

20. Автор поместил здесь текст письма Био (не указав его даты), *ясно изложившего принцип* указанного применения.

21. Это обстоятельство не было отмечено.

Библиография

Гаусс К. Ф. (1823, латин.), Теория комбинаций наблюдений и т. д. *Избр. геодезич. соч*., т. 1. М., 1957, с. 17 – 57.

Колмогоров А. Н. (1983), О логических основаниях теории вероятностей. В книге автора *Теория вероятностей и математическая статистика*. М., 1986, с. 467 – 471.

Buffon G. L. L. (1777), *Essai d’arithmétique morale*. *Oeuvr. Phil.* Paris, 1954, pp. 456 – 488.

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

Hume D. (1739), *Treatise on Human Nature*. Baltimore, 1969. [*Трактат о человеческой природе*. В *Сочинениях* автора (тт. 1 – 2. М., 1965).]

--- (1748), *An Enquiry concerning Human Understanding*. Переработанное автором сочинение 1739 – 1740 гг. Пуассон сослался на седьмой очерк французского издания этой переработки.

Lacroix C.-F. (1816), *Traité élémentaire du calcul des probabilités*. Paris, 1822, 1833, 1864.

Lamarck J. B. (1815), *Histoire naturelle des animaux sans vertèbres*, t. 1. Paris.

Poisson S.-D. (1824 – 1829), Sur la probabilité des résultats moyens des observations. *Conn. des Temps* pour 1827, pp. 273 – 302; pour 1832, pp. 3 – 22.

--- (1830), Sur la proportion des naissances des filles et des garçons. *Mém. Acad. Sci. Paris*, t. 9, pp. 239 – 308.

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

Глава 3. Исчисление вероятностей, зависящих от очень больших чисел

Опечатки/ошибки, не замеченные автором

1. В § 70, с . 179 оригинала. Формула в последней строке. Показатель степени первого сомножителя должен быть – $\mu/2$, а не $\mu/2$, см. следующую страницу оригинала.

2. В § 70, с. 181 оригинала. Пределы второго интеграла должны быть 0 и ∞, а не ∞ и x.

3. В § 72, с. 188 оригинала, через a' и b' были обозначены количества шариков остававшихся в урне, затем – количества извлечённых из неё. Далее, отношение p'/q' вначале было приравнено к a/b, затем – к a'/b'.

4. В § 74, с. 191 оригинала, выше формулы (10). Отношение двух интегралов: в правой части, в знаменателе, второй член ряда должен быть $(1 + \alpha)$, а не $(1 + 2)$.

5. В § 76, с. 195 оригинала. Последняя выделенная строка включает излишнее и странное равенство $K' = 1/2$.

6. В § 79, с. 202 оригинала, перед формулой для n/μ. Замечание автора по поводу выключенного в тексте выражения для n заканчивается, видимо, ошибочной ссылкой на формулу (6).

7. В самом конце § 86, с. 221 оригинала. В строке 2 над § 87 знак неравенства должен быть $\geq$, а не $\leq$.

8. В § 88, с. 225 оригинала, формула для v_1. Во втором слагаемом правой части перед корнем должна быть дробь $v\mu_1/\mu$, а не v/μ.

9. В § 90, с. 233 оригинала, выделенная формула после слов *Il en résultera.* Числитель в правой части содержит множитель $\varphi(a - g, b - g)$ вместо $\varphi(a - g, b - h)$.

10. В § 90, с. 234 оригинала. В конце параграфа автор заменяет произвольные величины a и b на $a + n$ и $b + m$, фактически же на $a + m$ и $b + n$.

11. В § 91, с. 234 оригинала. В первой строке числа $a, b, a - m, a - n$ следует заменить на $a, b, a - m, b - n$.

Все исправления внесены в переводе.

66. Если желательно вычислить отношение очень больших степеней двух данных чисел, всегда можно без труда применить таблицу логарифмов и при необходимости обратиться к логарифмам, содержащим больше десятичных знаков, чем в обычно применяемых […]. Но положение изменяется, если дело идёт об отношении двух произведений, каждое из которых состоит из очень большого числа неравных сомножителей […]. Приходится прибегать к приближённым методам, первый пример которых привёл Стирлинг […].

Эти отношения произведений очень большого числа сомножителей и сумм также очень большого числа отношений встречаются в большинстве важнейших приложений исчисления вероятностей. Поэтому правила, описанные в двух предыдущих главах, хоть они и были приведены в законченном виде, оказываются малополезными или вовсе обманчивыми, если не подкрепляются формулами для вычисления численных значений с достаточным приближением. […]

67. Прежде всего рассмотрим величину $n!$.[1] Интегрируя по частям, можно получить

$$\int_0^\infty e^{-x}x^n dx = n!. \qquad (67.1)$$

Подынтегральная функция исчезает при $x = 0$ и ∞. Внутри этих пределов она не становится бесконечной и проходит лишь через один максимум […]. Значение этого максимума обозначим через H, а через h – соответствующее значение x. Тогда $h = n$, $H = e^{-h}h^n$. Можно принять $e^{-x}x^n = H\exp(-t^2)$, где t – новая переменная, возрастающая от $-\infty$ до ∞. Её частные значения $t = -\infty, 0, \infty$ соответствуют значениям $x = 0, h, \infty$. […] Если $x = h + x'$, то […]

$$x' = h't + h''t^2 + \ldots$$

причём […]

$$1+\frac{1}{2}\frac{d^2\ln H}{dh^2}h'^2 = 0,\ \ \frac{d^2\ln H}{dh^2}h'' + \frac{1}{6}\frac{d^3\ln H}{dh^3}h'^2 = 0, \ldots \qquad (67.2)$$

[…]

$$n! = n^n e^{-n}\sqrt{2\pi n}\ \left(1+\frac{1}{12n}+\frac{1}{288n^2}+\ldots\right). \qquad (67.3)$$

68. Вначале ряд в формуле (67.3) сходится, тем скорее, чем больше n. Однако [следует фраза, исправленная самим автором], закон этого ряда неизвестен, он может относиться к тем, которые при достаточном продолжении становятся расходящимися[2]. Тем не менее, если ограничиться его сходящейся частью, всегда можно будет использовать его для вычисления [...], и для очень хорошего приближения даже не обязательно, чтобы n было весьма значительным.

[Следует пример при $n = 10$ и вывод формулы Валлиса.]

Анализ обязан Лапласу методом, который мы применяли для сведения интегралов в ряды, сходящиеся в своих первых членах, и который подходит для вычисления приближённых значений, если подынтегральные величины возводятся в очень большие степени. Ниже мы применим этот метод и в другом случае.

69. Далее, пусть E и F – противоположные события какой-либо природы, одно из которых наступает в каждом испытании. Их вероятности p и q мы предполагаем постоянными, и обозначим через U вероятность, что при μ испытаниях они появятся соответственно m и n раз. Тогда (§ 14)

$$U = C_{\mu}^{m} p^m q^n,\ m + n = \mu,\ p + q = 1. \qquad (69.1)$$

Если μ, m и n очень большие числа, для вычисления численного значения U следует применять формулу Стирлинга. А если указанные числа настолько велики, что в этой формуле можно пренебречь её последним сомножителем, то [...] и приближённо

$$U = \left(\frac{\mu p}{m}\right)^m \left(\frac{\mu q}{n}\right)^n \sqrt{\frac{\mu}{2\pi mn}}. \qquad (69.2)$$

Легко установить, что вероятнейшее составное событие, или то, для которого это значение U является наибольшим, соответствует наибольшему возможному приближению к равенству отношений m/n и p/q. Если, напротив, m и n – заданные числа, а p и q переменны, в сумме составляют единицу и могут возрастать бесконечно малыми приращениями от 0 до 1, максимум U по обычным правилам будет соответствовать значениям $p = m/\mu$ и $q = n/\mu$. Но ввиду многих других составных событий, [пусть] менее вероятных, его вероятность окажется незначительной и будет снижаться по мере дальнейшего возрастания μ, которое по предположению очень велико.

К примеру, при $p = q = 1/2$ и чётном μ вероятнейшее составное событие соответствует случаю $m = n = \mu/2$, и по формуле (69.2) его вероятность равна

$$U = \sqrt{2 / \pi\mu}. \qquad (69.3)$$

Видно, что она убывает обратно пропорционально $\sqrt{\mu}$. При $\mu = 100$, $U = 0.07979$, $1 - U = 0.92021$. Можно ставить немногим более 92 против восьми за то, что при 100 испытаниях равновероятные противоположные события E и F тем не менее не произойдут одно и то же число раз. Если же удерживать бесконечный ряд в формуле Стирлинга, то последнее выражение для U окажется умноженным на $(1 - 1/4\mu)$. […]

70. Составное событие, для которого m/n наиболее близко к p/q, является вероятнейшим, но при очень больших μ вероятности и других составных событий начинают быстро убывать при уклонении m/n от p/q в ту или иную сторону за определённые границы, протяжённостью обратно пропорциональные $\sqrt{\mu}$. Пусть снова $p = q = 1/2$ и g – заданная величина, положительная или отрицательная, меньшая без учёта знака, чем $\sqrt{\mu}$. Примем в формуле (69.2)

$$m = \frac{\mu}{2}[1 + \frac{g}{\sqrt{\mu}}],\ n = \frac{\mu}{2}[1 - \frac{g}{\sqrt{\mu}}], \qquad (70.1)$$

тогда

$$U = \left(1 - \frac{g^2}{\mu}\right)^{-\mu/2} \left(1 - \frac{g}{\sqrt{\mu}}\right)^{g\sqrt{\mu}/2} \left(1 + \frac{g}{\sqrt{\mu}}\right)^{-g\sqrt{\mu}/2} \sqrt{\frac{2}{\pi(\mu - g^2)}}.$$

Если g очень мало по сравнению с $\sqrt{\mu}$, то по биномиальной формуле (§ 8) три первых сомножителя в этом выражении для U будут почти точно равны

$\exp(g/^2 2)$, $\exp(- g/^2 2)$, $\exp(- g/^2 2)$.

Приняв под знаком радикала $\mu - g^2 = \mu$, мы получим

$$U = \sqrt{2/\pi\mu}\,\exp(-g^2/2)$$

как закон убывания U при небольшом удалении от своего максимума в ту или иную сторону. При $\mu = 200$ и $g = 1/\sqrt{2}$ вероятность событиям E и F, шансы которых равны друг другу, наступить 105 и 95 раз соответственно, относятся к вероятности их появления по 100 раз как $e^{-1/4}: 1 \approx 3:4$.

Формула (69.2) предполагает, что числа μ, m и n очень велики; если при этом m/n намного отклоняется от p/q, эта формула приведёт к весьма малому значению U относительно его максимума. Но уместно заметить, что при другом методе аппроксимирования при очень малой разности ($m/n - p/q$) по-прежнему очень малое найденное значение U может значительно отличаться от выводимого по формуле (69.2).

Чтобы прояснить это, я замечу, что, в соответствии с формулой в одном из моих мемуаров об определенных интегралах (1823, *J. Ecole Polyt.*, No. 19, p. 490),

$$\frac{2}{\pi}\int_0^{\pi/2}\cos^{\mu} x\cos[(m-n)x]dx \;=\; \frac{1}{2^{\mu}}C_{\mu}^{m}.$$

Каковы бы ни были числа m и n и их сумма μ, по формуле (69.1) в случае $p = q = 1/2$, которым мы ограничиваемся,

$$U \;=\; \frac{2}{\pi}\int_0^{\pi/2}\cos^{\mu} x\cos[(m-n)x]dx.$$

А если μ очень велико и в приближённых вычислениях рассматривается как бесконечное число, то при конечном x сомножитель $\cos^{\mu}x$ исчезает, другой же сомножитель, $\cos[(m - n)x]$, всегда конечен. Поэтому, не меняя значения интеграла, можно полагать, что его пределами являются лишь 0 и α при положительном и бесконечно малом α. В этих пределах

$$\cos x = 1 - x^2/2,\; \cos^{\mu} x = \exp(-\mu x^2/2),$$
$$U \;=\; \frac{2}{\pi}\int_0^{\alpha}\exp(-\mu x^2/2)\cos[(m-n)x]dx.$$

Фактически для любого конечного x экспоненциальный сомножитель исчезает, и, не изменяя значения этого нового интеграла, можно полагать его верхний предел бесконечным, так что по известной формуле

$$\int_0^{\infty}\exp(-\mu x^2/2)\cos[(m-n)x]dx = \sqrt{\frac{\pi}{2\mu}}\exp(-\frac{(m-n)^2}{2\mu}),$$

$$U = \sqrt{\frac{2}{\pi\mu}}\exp(-\frac{(m-n)^2}{2\mu}).$$

Примем теперь, как и выше [как следует из формулы (70.1)]

$$m - n = g\sqrt{\mu}.$$

Значение U совпадает с выводимым из выражения (69.2) только, если g очень мало по сравнению с $\sqrt{\mu}$. При других значениях g отношение указанных двух значений U будет весьма отличаться от единицы и может даже стать очень большим. Если, к примеру, $g = \sqrt{\mu}/2$, а $m - n = \mu/2$, то

$$U = \sqrt{2/\pi\mu}\, e^{-\mu/8}.$$

Из формулы (69.2) следует, что

$$U = (1-1/4)^{-\mu/2}(1-1/2)^{\mu/4}(1+1/2)^{-\mu/4}\sqrt{8/3\pi\mu},$$

а так как второй сомножитель здесь почти равен третьему, то

$$U = (9/8)^{-\mu/2}\sqrt{8/3\pi\mu}.$$

Значения U соответствуют друг другу в том смысле, что оба очень малы. Это доказывает, что при очень большом числе μ испытаний вероятность двум событиям, E и F, имеющим равные шансы появиться, наступить $3\mu/4$ и $\mu/4$ раз исключительно низка. Но отношение последнего значения U к первому

$$\frac{2}{\sqrt{3}}\left(\frac{64}{81}\sqrt{e}\right)^{\mu/4}$$

будет неограниченно возрастать с μ и при $\mu = 100$ уже превзойдёт 800.

71. Пусть, как и прежде, шансы E и F постоянны, но неизвестны. Установлено лишь, что при $\mu = m + n$ испытаниях они наступили m и n раз соответственно. Требуется определить вероятность U', что при $\mu_1 = m_1 + n_1$ новых испытаниях эти события произойдут m_1 и n_1 раз. По § 46

$$U' = C_{\mu_1}^{m_1} \frac{(m+m_1)!(n+n_1)!(\mu+1)!}{m!n!(\mu+\mu_1+1)!}.$$

[Следуют преобразования, приводящие к]

$$U' = C_{\mu_1}^{m_1} K(1+\frac{m_1}{m})^m (1+\frac{n_1}{n})^n (1+\frac{\mu_1}{\mu})^{-\mu} \times$$
$$(\frac{m+m_1}{\mu+\mu_1})^{m_1} (\frac{n+n_1}{\mu+\mu_1})^{n_1}, \qquad (71.1)$$

$$K = \frac{\mu+1}{\mu+\mu_1+1} \sqrt{\frac{(m+m_1)(n+n_1)(\mu+1)}{mn(\mu+\mu_1+1)}}. \qquad (71.2)$$

Если же m_1 и n_1 очень малы сравнительно с m и n, и если заменить K единицей, от которой оно очень мало отличается, то почти точно […]

$$U' = C_{\mu_1}^{m_1} (m/\mu)^{m_1} (n/\mu)^{n_1}.$$

Из формулы (69.1) следует, что последнее выражение совпадает с вероятностью того, что события E и F наступят m_1 и n_1 раз в $m_1 + n_1$ испытаниях, если их шансы p и q, известные заранее, достоверно равны $p = m/\mu$ и $q = n/\mu$. При $m_1 = 1$ и $n_1 = 0$ будет $C_{\mu_1}^{m_1} = 1$ и $U' = m/\mu$, что является вероятностью, что событие E, наступившее m раз при очень большом числе μ испытаний, появится вновь при следующем тираже, и это соответствует правилу § 49. Но когда m_1 и n_1 сравнимы по величине с m и n, указанное выше совпадение не будет иметь места.

Пусть для пояснения h – целое число или не очень малая дробь, $m_1 = mh$, $n_1 = nh$, $\mu_1 = \mu h$. Величина K будет почти равна $1/\sqrt{1+h}$, и ввиду равенства $\mu = m + n$ формула (71.1) сведётся к

$$U' = \frac{1}{\sqrt{1+h}} C_{\mu_1}^{m_1} (m/\mu)^{m_1} (n/\mu)^{n_1}.$$

Полагая $p = m/\mu$ и $q = n/\mu$, заменяя m и n на m_1 и n_1, обозначая результат через U_1 и сравнивая это выражение с формулой (69.1), мы получаем

$$U' = \frac{1}{\sqrt{1+h}} U_1.$$

Таким образом, U' меньше, чем U_1, и если h – очень большое число, оно окажется очень низкой вероятностью. Итак, разность между заданными вероятностями p и q событий E и F и их вероятностями m/μ и n/μ, полученными при очень большом числе испытаний, существенна. Именно, вероятность, что эти события появятся заданное число раз в новой серии испытаний, будет во втором случае ниже, чем в первом.

Различие вызвано тем, что вероятности E и F, выведенные из наблюдений, будь их сколь угодно много, всё же лишь вероятны, тогда как вероятности тех же событий, известные заранее, являются достоверными[3]. Если, к примеру, известно, что в урне находится равное число белых и чёрных шариков, то будет существовать вероятность, почти равная 0.07979, см. § 69, что в 100 тиражах с возвращением появятся по 50 шариков каждого цвета. Но если отношение шариков не задано, и известно лишь, что произошло указанное событие, то вероятность, что в 100 новых испытаниях повторится то же самое, при $h = 1$ в предыдущем уравнении будет равна лишь $0.07979/\sqrt{2} = 0.05658$.

72. Чтобы привести пример изменения шансов событий E и F во время испытаний, я предположу, что в урне находятся c шариков, a из них белые и b – чёрные. Без возвращения извлечено μ шариков, и пусть V будет вероятность, что в каком-то определённом порядке появилось m белых шариков и n чёрных. По формуле § 18, обозначив через a' и b' количества оставшихся в урне шариков, $a' + b' = c'$, мы получим в обозначениях § 71

$$V = C_{c'}^{a'} \frac{a!b!\mu!}{m!n!c!}.$$

Если a, b, m, n – очень большие числа, то при вычислении следует применять формулу Стирлинга. Имея в виду, что $\mu = m + n$ и $c = a + b$, мы заключаем, что приближённо

$$V = C_{c'}^{a'} \left(\frac{a}{c}\right)^{a} \left(\frac{b}{c}\right)^{b} \left(\frac{m}{\mu}\right)^{-m} \left(\frac{n}{\mu}\right)^{-n} \sqrt{\frac{ab\mu}{mnc'}}.$$

При $m = a$, $n = b$, $\mu = c$, так что $C_{c'}^{a'} = 1$, эта вероятность в точности равна единице, ибо после c испытаний из урны выйдут все, содержавшиеся в ней шарики, что соответствует достоверности.

Если $m/n = a/b$, то $a/c = m/\mu$ и $b/c = n/\mu$, и если $a/c = p'$ и $b/c = q'$, то

$$V = C_{c'}^{a'} p'^{a'} q'^{b'} \sqrt{c/\mu}.$$

Сравнивая это выражение с формулой (69.1) и обозначая через V' вероятность, что события с постоянными шансами, равными шансам a/c и b/c извлечения белого и чёрного шариков в начале испытаний, появляются a' и b' раз в c' испытаниях, мы имеем

$$V = V' \sqrt{c/\mu}. \tag{72.1}$$

Это доказывает, что $V/V' = \sqrt{c}/\sqrt{\mu}$ при любом числе c' шариков, остающихся в урне после этих тиражей, если только μ очень велико. Можно заметить, что $a' = p'(c - \mu)$, $b' = q'(c - \mu)$, так что отношение этих остающихся a' и b' шариков равно $p'/q' = a/b$. Если $p' = q' = 1/2$, так что $a' = b' = c'/2$, то (§ 69) $V' = \sqrt{2/\pi c'}$, а поскольку $c' = c - \mu$, $V = \sqrt{2c/[\pi\mu(c-\mu)]}$. Если $\mu = c/2$, то $V = \sqrt{4/\pi\mu} = V'\sqrt{2}$. Это означает, что если урна содержала поровну очень большое число белых и чёрных шариков, и из неё извлекли без возврата их половину, вероятность появления одного и того же числа шариков каждого цвета в $\sqrt{2}$ раз выше, чем если тиражи происходили бы с возвращением.

73. Я возвращаюсь к случаю постоянных шансов p и q событий E и F и рассматриваю вероятность, что после $\mu = m + n$ тиражей E появится не менее m раз, а F – не более n раз. Эта вероятность P будет суммой первых членов разложения $(p + q)^{\mu}$, расположенного по возрастающим степеням q. Однако, в указанном виде её трудно преобразовать в интеграл, чтобы при очень больших m и n можно было бы применить к нему метод § 67. Начнём с отыскания другого, более подходящего для нашей цели выражения для P. Можно также сказать, что интересующее нас составное событие G состоит в том, что F не появится более n раз в μ испытаниях. Эта цель будет достигнута в следующих (n + 1) случаях.

1) Если в каждом из m первых испытаний появилось событие E. Останется лишь $\mu - m = n$ испытаний, в течение которых F не может наступить более n раз. Вероятность этого случая равна p^m.

2) Если в (m + 1) первых испытаниях E произошло m раз и F – один раз, но не в первом тираже, иначе этот второй случай совпадёт с первым. Ясно, что в (n – 1) последующих испытаниях F не сможет наступить более (n – 1) раз и потому не появится более n раз во всех тиражах. Вероятность этого случая равна $p^m q$, если F произошло в определённом испытании, в противном же случае она равна $mp^m q$.

3) Если в первых (m + 2) тиражах E произошло m раз и F – дважды, но не при втором испытании, что необходимо и достаточно, чтобы этот случай не совпал ни с первым, ни со вторым. Соответствующая вероятность этого события, происшедшего в определённом порядке, равна $p^m q^2$, в противном же случае – $m(m + 1)p^m q^2/2$.

Таким образом можно, наконец, дойти до (n + 1)-го случая, в котором F не произойдёт в последнем испытании, чтобы этот случай не совпал ни с одним из предыдущих. Соответствующая вероятность будет равна

$$\frac{m(m+1)...(m+n-1)}{n!}p^m q^n.$$

Рассмотренные (n + 1) случаи отличаются друг от друга и представляют все различные существующие способы появления события G, и его полная вероятность будет равна сумме соответствующих вероятностей (§ 10):

$$P = p^m[1 + mq + \frac{m(m+1)}{2!}q^2 + ...$$
$$+ \frac{m(m+1)...(m+n-1)}{n!}q^n]. \quad (73.1)$$

Это выражение[4] должно совпадать с разложением $(p + q)^\mu$, но его легко преобразовать в определённые интегралы, чьи численные значения можно вычислить по методу § 67 с тем лучшим приближением, чем более велики числа m и n.

74. [Автор вычисляет интеграл

$$\int Xdx,\ X = \frac{x^n}{(1+x)^{\mu+1}},$$

интегрируя его по частям (n + 1) раз, и выводит формулу]

$$P = \frac{\int_\alpha^\infty Xdx}{\int_0^\infty Xdx} = \frac{1}{(1+\alpha)^m}[1 + m\frac{\alpha}{1+\alpha} + \frac{m(m+1)}{2!}\frac{\alpha^2}{(1+\alpha)^2} + ...$$

$$+ \frac{m(m+1)...(m+n-1)}{n!}\frac{\alpha^n}{(1+\alpha)^n}]. \qquad (74.1)$$

Здесь α – неотрицательное число. Если принять $\alpha = q/p$ и заметить, что $p + q = 1$, то правая часть последнего уравнения совпадёт с правой частью формулы (73.1). Для этого значения α левая часть указанного уравнения будет равна P и при $n = 0$ и $m = \mu$ она окажется вероятностью, что Е наступит не менее μ раз, т. е. в каждом испытании. Поэтому $P = p^\mu$. И действительно, при $n = 0$

$$\int_\alpha^\infty Xdx = \frac{1}{\mu(1+\alpha)^\mu} = \frac{1}{\mu}p^\mu,\quad \int_0^\infty Xdx = \frac{1}{\mu},\ P = p^\mu. \qquad (74.2)$$

При $n = \mu - 1$ и $m = 1$ вероятность P опишет наступление Е не менее одного раза, так что F не произойдёт во всех тиражах, и должна быть равной $P = 1 - q^m$, что можно проверить. Именно, пусть

$x = 1/y$, $dx = -dy/y^2$, $\alpha = 1/\beta$.

Для $n = \mu - 1$

$$\int_\alpha^\infty Xdx = \int_0^\beta \frac{dy}{(1+y)^{\mu+1}} = \frac{1}{\mu}[1 - \frac{1}{(1+\beta)^\mu}],\quad \int_0^\infty Xdx = \int_0^\infty \frac{dy}{(1+y)^{\mu+1}} = \frac{1}{\mu}$$

и ввиду $\beta = 1/\alpha = p/q$ и $1/(1 + \beta) = q$ формула (74.1) совпадёт с предыдущим выражением для P.

75. Применим вначале метод § 67 к интегралу (74.2). Как и в том параграфе, обозначим значение x, соответствующее максимуму X, через h, а сам этот максимум через H. Уравнение $dX/dx = 0$, из которого определяется h, будет иметь вид

$$n(1 + h) - (\mu + 1)h = 0$$

и потому

$$h = \frac{n}{m+1},\ H = \frac{n^n (m+1)^{m+1}}{(\mu+1)^{\mu+1}}. \qquad (75.1)$$

Если в уравнениях (67.2) принять

$$H = \frac{h^n}{(1+h)^{\mu+1}}$$

и после дифференцирований H по h подставить вместо этой переменной её предыдущее значение, то

$$h' = \sqrt{\frac{2(\mu+1)n}{(m+1)^3}},\ h'' = \frac{2(\mu+1+n)}{3(m+1)^2},\ldots$$

Если m, n и μ очень большие числа одного и того же порядка, то легко установить, что h', h'', … образуют очень быстро убывающий ряд, первый член которого, h', имеет порядок $1/\sqrt{\mu}$, второй, h'', – порядок $1/\mu$, третий – $1/(\mu\sqrt{\mu})$ и т. д., так что теперь

$$\int_0^\infty X dx = H\sqrt{\pi}(h' + \frac{1\cdot 3}{2}h''' + \frac{1\cdot 3\cdot 5}{4}h^{\mathrm{v}} + \ldots). \qquad (75.2)$$

76. Значение интеграла в числителе левой части формулы (74.1) зависит от того, будет ли $\alpha > h$ или $< h$, где h, как и раньше, является значением x, которое соответствует максимуму X. Переменная t в преобразовании в § 67 должна быть положительной при всех значениях $x > h$ и отрицательной при

$x < h$. Пусть, при $x = \alpha$, θ и A будут соответствовать значениям t и X. Тогда [следует вывод A и при $\alpha = q/p$; с учётом предыдущего значения H оказывается, что] $\theta = \pm k$, где

$$k^2 = n \ln \frac{n}{q(\mu+1)} + (m+1)\ln \frac{m+1}{p(\mu+1)}. \qquad (76.1)$$

Полагая k положительной величиной, следует принять $\theta = k$ при $q/p > h$ и $\theta = -k$ при $q/p < h$. Поэтому, в соответствии с преобразованием в § 67 мы имеем в первом случае

$$\int_\alpha^\infty X dx = H \int_k^\infty \exp(-t^2)\frac{dx'}{dt}dt.$$

Во втором случае мы получаем

$$H\int_{-k}^{\infty} \exp(-t^2)\frac{dx'}{dt}dt \ = \ H\int_{-\infty}^{\infty} \exp(-t^2)\frac{dx'}{dt}dt -$$
$$H\int_{-\infty}^{-k} \exp(-t^2)\frac{dx'}{dt}dt \qquad (76.2)$$

и, как в § 67,

$dx'/dt = h' + 2h''t + 3h'''t^2 + \ldots$

[…] Обозначим[5]

$$\int_k^\infty \exp(-t^2)t^{2i}dt = K_i, \ \int_k^\infty \exp(-t^2)t^{2i+1}dt = K_i'.$$

Поскольку первое слагаемое в формуле (76.2) равно интегралу (75.2), мы получаем при $q/p > h$, см. (75.1),

$$\int_\alpha^\infty X dx = H(h'K_0 + 3h'''K_1 + 5h^{\mathrm{V}}K_2 + \ldots) +$$
$$H(2h''K_0' + 4h^{\mathrm{IV}}K_1' + 6h^{\mathrm{VI}}K_2' + \ldots) = \mathrm{I} + \mathrm{II},$$
$$\int_\alpha^\infty X dx = \int_0^\infty X dx - \mathrm{I} + \mathrm{II} \text{ при } q/p < h. \qquad (76.3a, b)$$

Каждый ряд в этих формулах обладает, вообще говоря, той же степенью сходимости, что и (75.2). Если $k \neq 0$, значения интегралов K_i определяются только приближённо, но интегралы K'_i всегда можно представить в конечном виде

$$K'_i = (1/2)\exp(-k^2)(k^{2i} + ik^{2i-2} + i(i-1)k^{2i-4} + \ldots + i!k^2 + i!).$$

При $\alpha = h$ формулы (76.3a, b) должны будут совпадать. В то же время

$$\frac{q}{p} = \frac{n}{m+1},\ q = \frac{n}{\mu+1},\ p = \frac{m+1}{\mu+1}$$

и k, определяемое уравнением (76.1), исчезает. В результате

$$K_0 = \frac{\sqrt{\pi}}{2},\ K_i = 1\cdot 3\cdot 5\cdot \ldots \cdot (2i-1)\frac{\sqrt{\pi}}{2^{i+1}},\ K'_i = \frac{i!}{2}$$

и в соответствии с (75.2) формулы (76.3a, b) сводятся к единому выражению

$$\int_{\alpha}^{\infty} X dx = \frac{H\sqrt{\pi}}{2}\left(h' + \frac{1\cdot 3}{2}h''' + \frac{1\cdot 3\cdot 5}{4}h^{\mathrm{v}} + \ldots\right) + \\ H(h'' + 1\cdot 2h^{\mathrm{IV}} + 1\cdot 2\cdot 3h^{\mathrm{VI}} + \ldots).$$

77. Мы предположим, что числа m, n и μ достаточно велики, чтобы в различных формулах можно было бы пренебречь величинами h''', h^{IV} и т. д. Учитывая данные выше значения h' и h'', мы имеем

$$\frac{h''}{h'} = \frac{(\mu+1+n)\sqrt{2}}{3\sqrt{n(m+1)(\mu+1)}}$$

и при помощи уравнения (74.1) и формул (75.2) и (76.3) получаем соответственно при $q/p > h$ и $q/p < h$

$$P = \frac{1}{\sqrt{\pi}}\int_{k}^{\infty}\exp(-t^2)dt + \frac{(\mu+n)\sqrt{2}}{3\sqrt{\pi\mu mn}}\exp(-k^2), \qquad (77.1a)$$

$$P = 1 - \frac{1}{\sqrt{\pi}}\int_k^\infty \exp(-t^2)dt + \frac{(\mu+n)\sqrt{2}}{3\sqrt{\pi\mu mn}}\exp(-k^2). \quad (77.1b)$$

Здесь $k > 0$ и k^2 определяется по формуле (76.1).

Для упрощения мы подставили в последние члены правых частей этих формул μ и m вместо $\mu + 1$ и $m + 1$ и они с достаточным приближением определяют требуемое P. Если μ чётно и $m = n = \mu/2$ и $q > p$, то $h = \mu/(\mu + 2)$, $q/p > h$ и следует обратиться к формуле (77.1a). Учитывая уравнение (76.1), мы получаем

$$P = \frac{1}{\sqrt{\pi}}\int_k^\infty \exp(-t^2)dt + \sqrt{\frac{2}{\pi\mu}}\exp(-k^2), \qquad (77.2)$$

$$k^2 = \frac{\mu}{2}\ln\frac{\mu}{2q(\mu+1)} + \frac{\mu+2}{2}\ln\frac{\mu+2}{2p(\mu+1)}.$$

P выражает вероятность, что при очень большом чётном числе испытаний более вероятное событие F тем не менее не наступит чаще, чем противоположное событие E. Если обозначить через U вероятность, что оба они произойдут одно и то же число раз, то $P - U$ окажется вероятностью, что F наступит реже, чем E. При $p = q = 1/2$ очевидно, что $P - U$ будет также вероятностью противоположного, а $2P - U$ будет равно единице. Поэтому

$$U = \frac{2}{\sqrt{\pi}}\int_k^\infty \exp(-t^2)dt - 1 + 2\sqrt{\frac{2}{\pi\mu}}\exp(-k^2).$$

Это легко можно проверить. Ввиду [проведенных преобразований]

$$k^2 = \frac{1}{4\mu} + \frac{1}{4(\mu+2)} + \ldots$$

Сохраняя только члены порядка малости $1/\sqrt{\mu}$, мы имеем $k = 1/\sqrt{2\mu}$, $\exp(-k^2) = 1$,

$$\int_k^\infty \exp(-t^2)dt = \int_0^\infty \exp(-t^2)dt - \int_0^k \exp(-t^2)dt = \frac{\sqrt{\pi}}{2} - \frac{1}{\sqrt{2\mu}}$$

и $U = \sqrt{2/\pi\mu}$, что совпадает с формулой (69.3), которая была получена из формулы (69.2) при $m = n$ и $p = q$.

Если μ нечётно, и $m = (\mu - 1)/2$, и, как и раньше, $q > p$, то будет выполняться неравенство $q/p > h$. Из формул(77.2) и (76.1) следует, что

$$k^2 = \frac{\mu-1}{2}\ln\frac{\mu-1}{2q(\mu+1)} + \frac{\mu+3}{2}\ln\frac{\mu+3}{2p(\mu+1)}.$$

P – вероятность, что при очень большом числе μ испытаний более вероятное событие тем не менее произойдёт реже; при нечётном μ одинаково частое наступление Е и F невозможно. При $p = q = 1/2$ эта вероятность становится равной 1/2, что мы проверим. [Преобразуя выражение для k^2, автор получает]

$$k^2 = 1/(\mu-1) + 1/(\mu+3) + \ldots$$

Если пренебречь, как выше, членами порядка малости 1/μ, то

$$k = \frac{\sqrt{2}}{\mu},\ \exp(-k^2) = 1,\ \int_k^\infty \exp(-t^2)dt = \frac{\sqrt{\pi}}{2} - \frac{\sqrt{2}}{\mu}$$

и предыдущее значение P становится равным 1/2.

78. Пусть теперь n отличается от $(\mu + 1)q$ на положительную или отрицательную величину ρ, очень малую по сравнению с этим произведением. Поскольку $p + q = 1$ и $m + n = \mu$,

$$n = (\mu + 1)q - \rho,\ m + 1 = (\mu + 1)p + \rho.$$

Соответствующее значение h равно

$$h = \frac{(\mu+1)q-\rho}{(\mu+1)p+\rho} < \frac{q}{p} \text{ если мы примем } \rho > 0.$$

Раскладывая в ряд по степеням ρ правую часть уравнения (76.1), можно получить

$$k^2 = \frac{\rho^2}{2(\mu+1)pq}[1+\frac{(p-q)\rho}{3(\mu+1)pq}+...],$$

и, если ввести $r > 0$ по соотношению

$$\rho = r\sqrt{2(\mu+1)pq},\ \text{тогда}\ k \ = r[1+\frac{(p-q)r}{3\sqrt{2(\mu+1)pq}}+...],$$

и исключить случай очень малых p и q, полученный ряд в скобках будет быстро сходиться, поскольку он расположен по степеням $r/\sqrt{\mu+1}$ или $\rho/(\mu + 1)$. Сохраняя лишь первые два члена, можно для сокращения письма записать k в виде

$$k = r + \delta,\ \delta = \frac{(p-q)r^2}{3\sqrt{2(\mu+1)pq}}. \qquad (78.1)$$

В то же время

$$n = (\mu+1)q - r\sqrt{2(\mu+1)pq},$$

но во втором слагаемом в правой части формулы (77.1a) достаточно принять $k = r$ и заменить m и n на $p\mu$ и $q\mu$, чтобы вывести

$$P = \frac{1}{\sqrt{\pi}}\int_{r+\delta}^{\infty}\exp(-t^2)dt + \frac{(1+q)\sqrt{2}}{3\sqrt{\pi\mu pq}}\exp(-r^2). \qquad (78.2a)$$

Пусть [теперь] $\rho < 0$, тогда $h > q/p$. Обозначим через r' положительную величину и примем

$$\rho = -r'\sqrt{2(\mu+1)pq},$$

тогда

$$n = (\mu+1)q + r'\sqrt{2(\mu+1)pq}.$$

Но значение k, полученное из уравнения (76.1), должно неизменно оставаться положительным:

$$k = r' - \delta',\ \delta' = \frac{(p-q)r'^2}{3\sqrt{2(\mu+1)pq}}.$$

Формула (77.1b), которую мы должны применить, примет вид

$$P = 1 - \frac{1}{\sqrt{\pi}} \int\limits_{r'-\delta'}^{\infty} \exp(-t^2)dt + \frac{(1+q)\sqrt{2}}{3\sqrt{\pi\mu pq}} \exp(-r'^2). \qquad (78.2b)$$

Если вычесть из неё предыдущее значение P, то разность

$$R = 1 - \frac{1}{\sqrt{\pi}} \int\limits_{r+\delta}^{\infty} \exp(-t^2)dt - \frac{1}{\sqrt{\pi}} \int\limits_{r'-\delta'}^{\infty} \exp(-t^2)dt +$$
$$\frac{(1+q)\sqrt{2}}{3\sqrt{\pi\mu pq}} [\exp(-r'^2) - \exp(-r^2)] \qquad (78.3)$$

окажется вероятностью, что в очень большом числе μ испытаний количество наступлений события F не превзойдёт второго значения n и окажется большим, чем первое из них, по меньшей мере на единицу [конец фразы добавлен автором позже:] если вероятность выражается целым числом и не менее чем на единицу в противном случае.

79. Пусть для упрощения N – наибольшее целое число, содержащееся в μq и $f = \mu q - N$. Обозначим через u такую величину, чтобы $u\sqrt{2(\mu+1)pq}$ было целым числом, очень малым сравнительно с N. Примем также

$$q + f - r\sqrt{2(\mu+1)pq} = -u\sqrt{2(\mu+1)pq} - 1,$$
$$q + f + r'\sqrt{2(\mu+1)pq} = u\sqrt{2(\mu+1)pq}.$$

Границы значений n, относящиеся к вероятности R, станут равными

$$n = N - u\sqrt{2(\mu+1)pq} - 1,\ n = N + u\sqrt{2(\mu+1)pq}$$

и поэтому формула (78.3) выразит вероятность, что n по меньшей мере на единицу превзойдёт первую их них, но не окажется больше второй. Иначе говоря, выразит вероятность, что это число содержится в границах $N \pm u\sqrt{2\mu pq}$, равноудалённых от N (но μ было подставлено вместо μ + 1), или что оно совпадёт с какой-либо из них.

В соответствии с устанавливаемыми уравнениями и выражениями для δ и δ′

$$r + \delta = u + \varepsilon + \frac{1}{\sqrt{2(\mu+1)pq}}, \; r' - \delta' = u - \varepsilon,$$

где ε – величина, порядка малости 1/√μ. И если v – величина того же порядка, квадратом которой мы пренебрегаем, то

$$\int_{u+v}^{\infty} \exp(-t^2)dt = \int_{u}^{\infty} \exp(-t^2)dt - v\exp(-u^2).$$

Применяя это уравнение к обоим интегралам в формуле (78.3) и полагая $r' = r$ в членах вне знака интеграла, уже содержащих √μ в знаменателе, мы получаем

$$R = 1 - \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} \exp(-t^2)dt + \frac{1}{\sqrt{2\pi\mu pq}} \exp(-u^2), \qquad (79.1)$$

где в последнем члене μ заменило μ + 1.

Если желательно, чтобы интервал значений n не включал своей нижней границы, следует увеличить меньшее из значений n на единицу. В этом случае последнее слагаемое суммы r + δ, равно как и последний член формулы (79.1), исчезнут. Аналогично, чтобы указанный интервал не содержал своей верхней границы, следует уменьшить наибольшее значение n на единицу. Это уменьшает разность r' – δ′, и последний член формулы (79.1) также исчезнет. Наконец, знак этого члена следует изменить, чтобы интервал значений n не содержал ни одной из своих границ. Таким образом, этот член должен быть равен вероятности точного равенства

$$n = N + u\sqrt{2\mu pq},$$

где u – такая положительная или отрицательная величина, что второй член очень мал по сравнению с первым.

Если пренебречь величинами порядка малости $1/\mu$, то

$$\frac{n}{\mu} = q + u\sqrt{\frac{2pq}{\mu}},\ \frac{m}{\mu} = p - u\sqrt{\frac{2pq}{\mu}}$$

и [следуют преобразования] формула (69.2) принимает вид

$$U = \frac{\exp(-u^2)}{\sqrt{2\pi\mu pq}},$$

что и требовалось проверить.

Значение *P* в формуле (78.2a) является вероятностью, что *n* не превзойдёт границы $\mu q - r\sqrt{2\mu pq}$, где μ заменило μ + 1. Следовательно, если в значении *U* принять *u* = *r* и вычесть из него *P*, то разность *Q* = *P* − *U* окажется вероятностью того, что *n* не достигнет этой самой границы. Аналогично, если принять *u* = *r'* и вычесть *P*, соответствующее формуле (78.2b), то разность *Q'* = *P* − *U* окажется вероятностью того, что *n* будет ниже границы $\mu q + r'\sqrt{2\mu pq}$. Поэтому

$$Q = \frac{1}{\sqrt{\pi}} \int_{r+\delta}^{\infty} \exp(-t^2)dt + \frac{q-p}{3\sqrt{2\pi\mu pq}} \exp(-r^2), \qquad (79.2\text{a})$$

$$Q' = 1 - \frac{1}{\sqrt{\pi}} \int_{r'-\delta'}^{\infty} \exp(-t^2)dt + \frac{q-p}{3\sqrt{2\pi\mu pq}} \exp(-r'^2). \quad (79.2\text{b})$$

Вспомним, что здесь *r* и *r'* – положительные величины, очень малые по сравнению с √μ и что поэтому границы *n*, к которым относятся *Q* и *Q'*, мало отличаются в ту или иную сторону соответственно от μ*q*. В то же время значения δ и δ', содержащиеся в этих величинах, очень малы по сравнению с *r* и *r',* и если заменить μ + 1 на μ, то [, как в § 78,]

$$\delta = \frac{(p-q)r^2}{3\sqrt{2\mu pq}},\ \delta' = \frac{(p-q)r'^2}{3\sqrt{2\mu pq}}.$$

80. Разделив границы *n*, соответствующие формуле (79.1), на μ и учитывая то, что представляет *U*, мы получим границы

$$q-\frac{f}{\mu}\mp\sqrt{\frac{2pq}{\mu}}$$

отношения n/μ, соответствующая вероятность которых равна R. Если пренебречь дробью f/μ, то R, определяемое формулой (79.1), окажется вероятностью того, что разность $n/\mu - q$ находится в пределах $\pm u\sqrt{2pq/\mu}$. С изменёнными знаками они же окажутся границами разности $m/\mu - p$, потому что сумма этих разностей, $(m + n)/\mu - p - q = 0$.

Всегда можно принять u столь большим, что эта вероятность R станет сколь угодно мало отличаться от достоверности. И для этого даже не обязательно выбирать большое значение для u. Достаточно, к примеру, принять u, равное 4 или 5, чтобы интеграл $\int_u^\infty \exp(-t^2)dt$, а потому и $1 - R$, стали почти неощутимыми. Если величина u примет подобное значение и останется постоянной, границы разности $m/\mu - p$ будут сжиматься по мере того, как μ, уже предположенное очень большим, станет ещё больше возрастать. Отношение m/μ числа появлений Е к общему числу испытаний будет отличаться от шанса p этого события все менее и менее. Число μ всегда можно достаточно увеличить, чтобы вероятность R разности $m/\mu - p$ оказалась сколь угодно низкой. Обратно, если неизменно увеличивать μ и принять за каждую указанную границу постоянную и заданную величину l, т. е., если обеспечить возрастание u в том же отношении, что и $\sqrt{\mu}$, значение R будет неограниченно приближаться к единице. Всегда можно достаточно увеличить μ, чтобы вероятность R того, что $m/\mu - p$ окажется в границах $\pm\, l$, сколь угодно мало отличалась от достоверности. В этом состоит теорема Якоба Бернулли (§ 49).

81. В предыдущем вычислении мы (§ 78) исключили случай очень малых шансов p или q. Пусть теперь q – очень малая дробь, т. е. вероятность события F очень низка. При очень большом числе μ испытаний отношение n/μ количества наступлений F к μ также будет очень малой дробью. Приняв $\mu - n$ вместо m в формуле (73.1), мы получим $q\mu = w$, $q = w/\mu$. Если пренебречь дробью n/μ, правая часть этой формулы примет вид

$$p^m\left(1+w+\frac{w^2}{2!}+\frac{w^3}{3!}+...+\frac{w^n}{n!}\right).$$

В то же время

$$p = 1 - \frac{w}{\mu},\ p^m = \left(1 - \frac{w}{\mu}\right)^{\mu} \left(1 - \frac{w}{\mu}\right)^{-n}.$$

Можно заменить первый сомножитель экспоненциальной функцией e^{-w}, а второй сомножитель – единицей. В соответствии с уравнением (73.1) мы получим тогда почти точно

$$P = (1 + w + \frac{w^2}{2!} + \frac{w^3}{3!} + ... + \frac{w^n}{n!})e^{-w}$$

для вероятности, что событие, шанс которого в каждом испытании равен очень малой дроби w/μ, не появится более n раз при очень большом числе μ испытаний.

При $n = 0$ это значение P становится равным e^{-w}. Такова, стало быть, вероятность, что рассматриваемое событие не наступит ни одного раза при μ испытаниях, и, следовательно, вероятность $1 - e^{-w}$, что оно произойдёт по крайней мере один раз, что уже было отмечено в § 8. Если же n не будет очень небольшим, значение P очень мало отклонится от единицы. Это можно усмотреть, записав P в виде

$$P = 1 - \frac{w^{n+1}e^{-w}}{(n+1)!}[1 + \frac{w}{n+2} + \frac{w^2}{(n+2)(n+3)} + ...].$$

Пусть $w = 1$ и $n = 10$, тогда разность $1 - P$ окажется почти точно равной $1/10^8$. Иными словами, почти достоверно, что событие с очень слабым шансом $1/\mu$ появления в каждом из μ испытаний не наступит более 10 раз.

82. Интеграл в формуле (79.1) вычисляется, вообще говоря, в квадратурах. В конце книги Kramp (1798) помещена таблица его значений с аргументом $u = 0$ до $u = 3$, в соответствии с которой

$$\int_3^{\infty} \exp(-t^2)dt = 0.00001957729...$$

Интегрируя по частям при $u > 3$ можно получить [автор вывел формулу

$$\int_0^u \exp(-t^2)dt = u - \frac{u^3}{1\cdot 3} + \frac{u^5}{1\cdot 2\cdot 5} - \frac{u^7}{1\cdot 2\cdot 3\cdot 7} + ...$$

и заметил, что ряд в правой части быстро сходится для значений u меньше единицы]. Если желательно вычислить значение u, при котором $R = 1/2$, то можно приравнять этот ряд к

$$\frac{\sqrt{\pi}}{4} - \frac{\exp(-u^2)}{2\sqrt{2\mu pq}}.$$

Пусть $u = a$ будет корнем полученного таким образом уравнения. Пренебрегая вторым слагаемым в правой части, мы можем записать с точностью $1/\mu$, что

$$u = a - \frac{1}{2\sqrt{2\mu pq}}.$$

После нескольких попыток оказалось, что $a \approx 0.4765$. Это означает, что разность $m/\mu - p$ с равной вероятностью будет заключена либо в пределах

$$\pm\left(0.4765\sqrt{\frac{2pq}{\mu}} - \frac{1}{2\mu}\right),$$

либо вне их. Для какого-нибудь значения u существует вероятность R, что разность между $m/\mu - p$ и $n/\mu - q$ окажется не большей, чем $\pm 2u\sqrt{2pq/\mu}$. Поэтому, если $p = q = 1/2$, то с вероятностью 1/2 величина $(m - n)/\mu$ окажется в пределах

$$\pm\left[\frac{0.6739}{\sqrt{\mu}} - \frac{1}{\mu}\right].$$

Таким образом, если шансы E и F одни и те же, разности количеств их появления с равной вероятностью будут, без учёта знака, больше или меньше $0.6739\sqrt{\mu} - 1$.

Пусть A и B играют очень много партий, например 10^6, справедливой игры. Тогда можно держать пари на равных, что кто-то из них выиграет на 674 партии больше другого. Вот эта разность, которая может равным образом благоприятствовать каждому, и является долей случая. Но если в каждой партии шанс p игрока А превышает шанс q игрока В, появится вероятность R, неизменно повышающаяся с μ, что А выиграет на $\mu(p - q) \pm 2u\sqrt{2\mu pq}$ больше, чем В. А так как первое слагаемое, появившееся ввиду разности умения игроков, возрастает пропорционально числу партий, тогда как второе возрастает лишь пропорционально корню квадратному из этого числа, то более умелый игрок в конце концов всегда окажется победителем, какой бы малой ни была разность $p - q$.

83. Предполагая, что шансы p и q были известны, мы определяли с высокой вероятностью и хорошим приближением отношения m/μ и n/μ при очень больших μ. Обратно, если шансы заранее неизвестны, но получены указанные отношения, выведенные нами формулы с хорошим приближением устанавливали весьма вероятные значения p и q. Именно, существует вероятность R, вычисляемая по формуле (79.1), что шанс p события Е заключён в границах $m/\mu \pm u\sqrt{2pq/\mu}$. Если R очень мало отличается от единицы, дробь p будет почти точно и с высокой вероятностью равна m/μ, а q равно n/μ. Подставляя поэтому m/μ и n/μ вместо p и q в двузначные члены их границ и в последнее слагаемое формулы (79.1), которое уже делилось на $\sqrt{\mu}$, мы получим вероятность

$$R = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt + \sqrt{\frac{\mu}{2\pi mn}}\exp(-u^2) \qquad (83.1)$$

того, что p заключено в пределы

$$\frac{m}{\mu} \pm \frac{u}{\mu}\sqrt{\frac{2mn}{\mu}}.$$

Если m, n и μ очень большие числа, можно, вообще говоря, применить приближённые значения m/μ и n/μ шансов p и q для вычисления вероятности будущего события, составленного из Е и F, к примеру, их появления m' и n' раз в $\mu' = m' + n'$ новых испытаниях, если μ' очень мало по сравнению с μ. Формулу (79.1) можно применять даже если μ' очень большое число. Именно, если подставить μ', m/μ и n/μ вместо μ, p, q в саму формулу и в соответствующие границы, то

$$R = 1 - \frac{2}{\sqrt{\pi}} \int_u^\infty \exp(-t^2)dt + \frac{\mu}{\sqrt{2\pi\mu' mn}} \exp(-u^2) \qquad (83.2)$$

выразит вероятность того, что n' заключено в границах

$$\frac{\mu' n}{\mu} \mp \frac{u}{\mu} \sqrt{2\mu' mn}\,. \qquad (83.3)$$

Здесь $\mu' n/\mu$ внесено вместо наибольшего целого числа, содержавшегося в этом отношении. Как бы близки ни были m/μ и n/μ к p и q, они лишь вероятны, а не достоверны, и их нельзя применять, как было замечено в § 71, когда μ' сравнимо с μ. Вот почему мы иначе рассмотрим вопрос об установлении p и q по наблюдённым событиям, чтобы применить их к вероятностям будущих событий.

84. Пусть, как и прежде, наступление событий E и F наблюдалось m и n раз в очень большом числе $\mu = m + n$ испытаний, в течение которых шансы этих событий p и q не изменялись. По предыдущему, существует весьма высокая вероятность, что эти шансы очень немного отличаются от отношений m/μ и n/μ, которые могут быть приняты за их приближённые значения. Эти шансы могут принимать бесконечное число значений, возрастающих бесконечно малыми приращениями, так что вероятности какого-то точного значения p и соответствующего значения q бесконечно малы. Их, или по крайней мере тех, которые мало отклоняются от m/μ и n/μ, и которые нам только и требуются, мы должны определить.

Величина Q, определяемая формулой (79.2а), это вероятность того, что $n < \mu q - r\sqrt{2\mu pq}$ и одновременно того, что неизвестный шанс события F, появившегося n раз в μ испытаниях, $q > n/\mu + r\sqrt{2pq/\mu} = n/\mu + (r/\mu)\sqrt{2mn/\mu}$, если вместо p и q подставить их приближённые значения m/μ и n/μ. Далее, если в эту формулу подставить $r - dr$ вместо r и сохранить лишь бесконечно малые первого порядка, то $Q - dQ/dr$ также станет вероятностью, что

$$q > \frac{n}{\mu} + \frac{r}{\mu}\sqrt{\frac{2mn}{\mu}} - \sqrt{\frac{2mn}{\mu}}\frac{dr}{\mu}.$$

Поэтому – dQ/dr выразит бесконечно малую вероятность, что в точности

$$q = n/\mu + (r/\mu)\sqrt{2mn/\mu}$$

при всех положительных значениях r, очень малых относительно $\sqrt{\mu}$, как это и предполагалось в формуле для Q. Аналогично, формула (79.2b) выразит вероятность Q' того, что $q > n/\mu - (r'/\mu)\sqrt{2mn/\mu}$. Подставляя $r' + dr'$ вместо r', мы получаем $Q' + (dQ'/dr')dr'$ для вероятности того, что

$$q > \frac{n}{\mu} - \frac{r'}{\mu}\sqrt{\frac{2mn}{\mu}} - \sqrt{\frac{2mn}{\mu}}\frac{dr'}{\mu}.$$

Следовательно, $(dQ'/dr')dr'$ будет вероятностью, что q превышает второй, но не первый предел, или что в точности

$$q = n/\mu - (r'/\mu)\sqrt{2mn/\mu}.$$

Здесь r' также является положительной величиной, очень малой по сравнению с $\sqrt{\mu}$.

[Автор преобразует выражения для dQ/dr и dQ'/dr' с точностью до малых порядка $1/\sqrt{\mu}$ и замечает, что они переходят одно в другое при перестановке r и $-r'$. Он обозначает] через v положительную или отрицательную переменную, очень малую сравнительно с $\sqrt{\mu}$ [и вводит]

$$V = \frac{1}{\sqrt{\pi}}\exp(-v^2) - \frac{2(m-n)v^2}{3\sqrt{2\pi\mu mn}}\exp(-v^2), \qquad (84.1)$$

причём Vdv оказывается вероятностью того, что

$$q = n/\mu + (v/\mu)\sqrt{2mn/\mu}.$$

Указанная бесконечно малая вероятность одновременно относится к

$$p = m/\mu - (v/\mu)\sqrt{2mn/\mu}\,. \qquad (84.2)$$

Величина V, как заметно, очень быстро убывает по мере возрастания v. И до того, как порядок переменной v сравняется с $\sqrt{\mu}$, порядок V окажется исключительно малым ввиду сомножителя exp $(- v^2)$. Если таким же образом выразить при помощи v значения p и q, весьма уклоняющиеся от m/μ и n/μ, и представить их вероятности через $V'dv$, где V' – функция v, отличная от V с намного меньшими численными значениями, чем у неё, то можно будет считать величину V' совсем неощутимой, что освобождает нас от обязанности отыскивать её.

Итак, пусть E′ – будущее событие, составленное из E и F, и П – вероятность, что оно произойдёт при определённых значениях шансов E и F. Таким образом, П – заданная функция p и q. Далее, пусть П′ – истинная вероятность E′ с учётом значений p и q, подставленных в П. При умножении П на эту бесконечно малую вероятность p и q и интегрировании произведения от $p = 0$ и $q = 1$ до $p = 1$ и $q = 0$ мы получим П′. Но, с учётом дальнейшего, можно пренебречь частью этого интеграла, соответствующей значительным уклонениям p и q от m/μ и n/μ. Поэтому, если подставить в П предыдущие значения p и q, то

$$\Pi' = \int \Pi V dv\,, \qquad (84.3)$$

где интегрирование происходит по положительным и отрицательным значениям v, очень малым относительно $\sqrt{\mu}$. Этот результат соответствует полученным мной более непосредственно (1830, § 2).

85. Для первого применения формул (84.1) и (84.3) я предположу, что П′ – вероятность, что при очень большом числе $\mu' = m' + n'$ новых испытаний события E и F наступили m' и n' раз, причём отношение m'/n' очень близко к отношению m/n количеств появления этих событий в μ прежних испытаниях. Иначе говоря, я предположу, что

$$m' = mh - \alpha\sqrt{\mu'},\; n' = nh + \alpha\sqrt{\mu'},\; \mu' = \mu h,$$

где h и α заданы, причём α может быть положительным и отрицательным, но очень малым относительно $\sqrt{\mu'}$.

По формуле (69.2) при

$$U' = \sqrt{\mu'/2\pi m'n'}$$

окажется, что

$$\Pi = U'\left(\frac{\mu' p}{m'}\right)^{m'}\left(\frac{\mu' q}{n'}\right)^{n'}$$

и можно уже заметить, что U' – вероятность события E′, если полагать, что m/μ и n/μ являются точными и заданными шансами E и F и при α = 0. Кроме того, ввиду того, что

$$\frac{m'}{\mu'} = \frac{m}{\mu} - \frac{\alpha}{\sqrt{\mu'}},\quad \frac{n'}{\mu'} = \frac{n}{\mu} + \frac{\alpha}{\sqrt{\mu'}}\,,$$

обозначая

$$\frac{v}{\mu}\sqrt{\frac{2mn}{\mu}} - \frac{\alpha}{\sqrt{\mu'}} = v_1,$$

можно записать значения p и q из § 84 в виде

$$p = (m'/\mu') - v_1,\ q = (n'/\mu') + v_1.$$

Подставляя эти значения в Π, мы получим

$$\Pi = U'\left(1 - \frac{\mu' v_1}{m'}\right)^{m'}\left(1 + \frac{\mu' v_1}{n'}\right)^{n'}.$$

Поскольку эти дроби имеют порядок $1/\sqrt{\mu}$ или $1/\sqrt{\mu'}$, то […]

$$\Pi = U'\exp\left(-\frac{\mu'^3 v_1^2}{2m'n'}\right). \qquad (85.1)$$

По той же причине можно пренебречь вторым слагаемым в формуле (84.1), так что [после преобразований] формула (84.3) примет вид

$$\Pi' = \frac{1}{\sqrt{\pi}} U' \int \exp(-v^2 - \frac{\mu'^3 v_1^{\,2}}{2m'n'}) dv \,.$$

Этот интеграл должен распространяться лишь на значения v, очень малые относительно $\sqrt{\mu}$. Если заметить, что ввиду экспоненциального сомножителя коэффициент dv при значениях v, сравнимых с $\sqrt{\mu}$, становится совсем неощутимым, можно распространить интеграл на подобные значения v и принять, как мы и будем делать, за его пределы $-\infty$ и ∞. Подставляя также mh и nh вместо m' и n' в значение v_1, мы получим

$$v^2 + \frac{\mu'^3 v_1^{\,2}}{2m'n'} = v^2(1+h) - \frac{2v\alpha\mu'\sqrt{h}}{\sqrt{2m'n'}} + \frac{\alpha^2\mu'^2}{2m'n'}$$

и примем

$$v\sqrt{1+h} - \frac{\alpha\mu'\sqrt{h}}{\sqrt{2m'n'(1+h)}} = x, \; dv = \frac{dx}{\sqrt{1+h}} \,.$$

Пределы интеграла при этой новой переменной x всё ещё будут бесконечными, так что искомая вероятность окажется равной

$$\Pi' = \frac{1}{\sqrt{1+h}} U' \exp[-\frac{\alpha^2\mu'^2}{2m'n'(1+h)}] \qquad (85.2)$$

с очевидным упрощением при $\alpha = 0$. Ввиду смысла U' это совпадает с результатом, полученным иным способом в § 71.

86. Для второго примера приложения формул (84.1) и (84.3) мы предположим, что Π' – это вероятность того, что разность $n'/\mu' - n/\mu$ не превзойдёт величины $\alpha/\sqrt{\mu'}$, которую она должна была достичь в предыдущем примере § 85.

Величина Π будет функцией шансов p и q событий E и F, а именно вероятностью, что в μ' будущих испытаниях F не наступит более n' раз, $n' = (n\mu'/\mu) + \alpha\sqrt{\mu'}$, а E произойдёт m' раз, по меньшей мере равным $(m\mu'/\mu) - \alpha\sqrt{\mu'}$, ср. формулу (85.1).

Значение этой вероятности будет выражено одной из формул (77.1) после подстановки μ', m', n' вместо μ, m, n.

Если неизменно ограничиваться величинами порядка $1/\sqrt{\mu}$ или $1/\sqrt{\mu'}$, то для указанных крайних значений m' и n' будет иметь место равенство

$$\frac{n'}{m'+1} = \frac{n}{m}[1 + \frac{\alpha\mu^2}{mn\sqrt{\mu'}}].$$

В соответствии со значениями p и q из § 85 в то же время будет выполняться соотношение

$$\frac{q}{p} = \frac{n}{m}[1 + v\sqrt{\frac{2\mu}{mn}}].$$

Если же переменная v ограничена без учёта знака следующим образом

$$v < \frac{\alpha\mu^2}{\sqrt{2\mu\mu' mn}}, \text{ то } \frac{q}{p} < \frac{n'}{m'+1} \text{ или } \frac{q}{p} > \frac{n'}{m'+1}$$

для положительных и отрицательных значений константы α. В этих случаях, ввиду формул (77.1b) и (77.1a) соответственно, оказывается, что

$$\Pi = 1 - \frac{1}{\sqrt{\pi}}\int_k^\infty \exp(-t^2)dt + \frac{2(\mu'+n')}{3\sqrt{2\pi\mu' m' n'}}\exp(-k^2),$$

$$\Pi = \frac{1}{\sqrt{\pi}}\int_k^\infty \exp(-t^2)dt + \frac{2(\mu'+n')}{3\sqrt{2\pi\mu' m' n'}}\exp(-k^2).$$

Величина $k > 0$, а её квадрат определяется формулой (76.1) при замене m, n, μ на m', n', μ'. Здесь должны быть применены крайние значения m' и n', равно как и величин p и q. В результате

$$q = \frac{n'}{\mu'+1} - v',\ p = \frac{m'+1}{\mu'+1} + v',$$

где для сокращения записи принято, что

$$\frac{\alpha}{\sqrt{\mu'}}-\frac{v\sqrt{2mn}}{\mu\sqrt{\mu}}-\frac{n'}{\mu'(\mu'+1)} = v'.$$

Величина v' имеет порядок $1/\sqrt{\mu'}$. [После преобразований автор получает]

$$k^2 = \frac{\mu'^3 v'^2}{2m'n'}-\frac{(m'-n')\mu'^4 v'^3}{3m'^2 n'^2},\ \ k=\pm k'[1-\frac{2(m'-n')k'}{3\sqrt{2\mu' m'n'}}],$$

$$k'=\frac{\alpha\mu'}{\sqrt{2m'n'}}-\frac{v\mu'\sqrt{\mu' mn}}{\mu\sqrt{\mu m'n'}}. \qquad (86.1)$$

Ввиду границы, назначенной для v, величина k' имеет тот же знак, что и α. Чтобы k было положительно при $\alpha > 0$ и $\alpha < 0$ требуется принять верхний или нижний знак в его выражении. [После преобразований, и, в частности, вывода новых формул для П, автор продолжает:] Учитывая формулы (84.1) и (84.3), соответствующие выражения для П′ оказываются такими:

$$\Pi'=\frac{1}{\sqrt{\pi}}\int\exp(-v^2)dv-\frac{1}{\pi}\int_{k'}^{\infty}\int\exp(-t^2-v^2)dtdv+$$
$$\frac{2n'}{\sqrt{2\pi\mu' m'n'}}\int\exp(-k'^2-v^2)dv-\ \frac{2(m-n)}{3\sqrt{2\pi\mu mn}}[\int\exp(-v^2)v^3dv-$$
$$\frac{1}{\sqrt{\pi}}\int_{k'}^{\infty}\int\exp(-t^2-v^2)v^3dtdv], \qquad (86.2a)$$

$$\Pi'=\frac{1}{\pi}\int_{-k'}^{\infty}\int\exp(-t^2-v^2)dtdv+\frac{2n'}{\sqrt{2\pi\mu' m'n'}}\int\exp(-k'^2-v^2)dv-$$
$$\frac{2(m-n)}{3\pi\sqrt{2\mu mn}}\int_{-k'}^{\infty}\int\exp(-t^2-v^2)v^3dtdv]. \qquad (86.2b)$$

[Автор упрощает эти формулы, распространяя, в частности, границы v на всю бесконечную прямую и вводит]

$$\frac{\alpha\mu'}{\sqrt{2m'n'}}=\pm\,\beta,\ \frac{\mu'\sqrt{\mu' mn}}{\mu\sqrt{\mu m'n'}}=\gamma,\ t=\theta\pm\gamma v,\ dt=d\theta.$$

Здесь $\beta > 0$ а знаки плюс и минус соответствуют знакам α. В формулах (86.2a, b) соответственно $\alpha > 0$ и < 0 и

$$k' = \pm\beta - \gamma v,\ t = \theta \mp \gamma v$$

и пределы интегралов по θ в обоих случаях всё ещё равны β и ∞. Поэтому

$$\Pi' = 1 - \frac{1}{\pi}\int_{\beta}^{\infty}\int_{-\infty}^{\infty}\exp[-\theta^2 + 2\gamma\theta v - (1+\gamma^2)v^2]d\theta dv +$$
$$\frac{2n'}{\sqrt{2\pi\mu' m' n'}}\int_{-\infty}^{\infty}\exp[-\beta^2 + 2\gamma\beta v - (1+\gamma^2)v^2]dv +$$
$$\frac{2(m-n)}{3\pi\sqrt{2\mu mn}}\int_{\beta}^{\infty}\int_{-\infty}^{\infty}\exp[-\theta^2 + 2\gamma\theta v - (1+\gamma^2)v^2]v^3 d\theta dv,$$

$$\Pi' = \frac{1}{\pi}\int_{\beta}^{\infty}\int_{-\infty}^{\infty}\exp[-\theta^2 - 2\gamma\theta v - (1+\gamma^2)v^2]d\theta dv +$$
$$\frac{2n'}{\sqrt{2\pi\mu' m' n'}}\int_{-\infty}^{\infty}\exp[-\beta^2 - 2\gamma\beta v - (1+\gamma^2)v^2]dv -$$
$$\frac{2(m-n)}{3\pi\sqrt{2\mu mn}}\int_{\beta}^{\infty}\int_{-\infty}^{\infty}\exp[-\theta^2 - 2\gamma\theta v - (1+\gamma^2)v^2]v^3 d\theta dv.$$

[…] Первое значение Π' – это вероятность того, что

$$n' \le \frac{n\mu'}{\mu} + \beta\sqrt{\frac{2m'n'}{\mu'}},$$

причём сумма в правой части очень немного превышает своё первое слагаемое. Второе значение Π' – это вероятность, что

$$n' \ge \frac{n\mu'}{\mu} - \beta\sqrt{\frac{2m'n'}{\mu'}},$$

причём разность в правой части немного меньше своего первого слагаемого.

87. Можно заметить, что ввиду бесконечных пределов по v первые два интеграла выражений для Π' совпадают, третьи же интегралы равны по величине, но противоположны по знаку. Первое выражение превосходит второе [после преобразований] на

$$\varphi = 1 - \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} \exp(-t^2)dt,$$

где $u = \beta/\sqrt{1+\gamma^2}$. Учитывая значение γ, можно заметить, что φ выражает вероятность, что n' заключено в границах

$$\frac{n\mu'}{\mu} \mp \frac{u\sqrt{2(\mu^3 m'n' + \mu'^3 mn)}}{\mu\sqrt{\mu\mu'}} \qquad (87.1)$$

или равно верхней границе. Если желательно, чтобы интервал между этими границами включил и нижнюю границу, следует добавить со знаком плюс к φ вероятность, что n' в точности равно ей, см. формулу (85.2), полагая $\alpha\sqrt{\mu'}$ равным второму слагаемому выражения (87.1)[6].

И если обозначить через w вероятность, что n' окажется внутри указанных границ, или равно какой-либо из них, то

$$w = 1 - \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} \exp(-t^2)dt +$$

$$\frac{\sqrt{\mu\mu'}}{\sqrt{2\pi m'n'(\mu+\mu')}} \exp[-\frac{u^2(\mu^3 m'n' + \mu'^3 mn)}{\mu^2 m'n'(\mu+\mu')}]. \qquad (87.2)$$

При сравнении этого значения w с R в формуле (83.2) видно, что они отличаются только своими последними слагаемыми и поэтому почти равны друг другу. Но если число μ' будущих испытаний не очень мало относительно μ, вторые слагаемые в выражениях (87.1) и (83.3) не совпадут для границ n', которые соответствуют w и R, и границы, вероятность которых равна w, могут оказаться намного менее тесными, чем во втором случае.

Если w мало отличается от достоверности, то в соответствующих границах можно вместо n' и m' подставить их приближённые значения, $n\mu'/\mu$ и $m\mu'/\mu$. Границы примут вид

$$\frac{n\mu'}{\mu} \mp \frac{u}{\mu}\sqrt{2\mu' mn(1+h)},$$

где $h = \mu'/\mu$. При сравнении этих пределов с пределами (83.3) видно, что при тех же значениях u границы расширились в $\sqrt{1+h}$ раз, а чтобы они сравнялись, значение u придётся уменьшить в $1/\sqrt{1+h}$ раз, что понизит их вероятность относительно R. Если h – очень небольшая дробь, формулы (83.2) и (87.2) почти совпадут, равно как и соответствующие границы значений n'. Этот результат согласуется с найденным мной иным методом (1830).

Формула (87.2) выражает также вероятность, что для разности $(n'/\mu' - n/\mu)$ имеют место границы

$$\mp \frac{u\sqrt{2(\mu^3 m'n' + \mu'^3 mn)}}{\mu\mu'\sqrt{\mu\mu'}} \tag{87.3}$$

и что они же с противоположными знаками имеют место и для $(m'/\mu' - m/\mu)$. И если принять u равным трём или четырём, так что вероятность w станет очень близка достоверности (§ 80)[7], и если, тем не менее, упомянутые выше наблюдённые разности заметно выходят за свои границы, можно будет обоснованно с очень высокой вероятностью заключить, что неизвестные шансы событий E и F изменились либо во время испытаний, либо между их сериями.

Можно заметить, что при одном и том же значении u и, следовательно, с той же самой вероятностью указанные выше границы окажутся шире всего при $\mu = \mu'$ и более тесными, когда одно из этих чисел очень велико по сравнению с другим. При $\mu' = \mu$ почти равны будут и m' и m и n' и n. При этом коэффициент при u [в выражении (87.3)] станет равным $2\sqrt{mn}/\mu\sqrt{\mu}$. Если, напротив, μ' очень велико по сравнению с μ, то почти точно $m' = m\mu'/\mu$ и $n' = n\mu'/\mu$ и этот коэффициент станет равным $\sqrt{2mn}/\mu\sqrt{\mu}$, т. е. в $\sqrt{2}$ раз меньше.

88. Пусть противоположные события E и F с неизвестными шансами p и q произошли m и n раз в очень большом числе μ испытаний. Другие противоположные события E_1 и F_1 также с неизвестными шансами p_1 и q_1 произошли m_1 и n_1 раз в очень большом числе μ_1 испытаний. Если отношения m/μ и m_1/μ_1, равно как и n/μ и n_1/μ_1, значительно отличаются друг от друга, следует считать достоверным или почти достоверным, что $p \neq p_1$ и $q \neq q_1$. Но когда разности указанных отношений являются небольшими дробями, возможно, что неравенства соответствующих шансов не являются ощутимыми и что они произошли потому, что события вовсе не наступали в точности пропорционально своим шансам. Полезно поэтому определить, что мы и сделаем, вероятность неравенств $p \neq p_1$ и $q \neq q_1$, соответствующую известным небольшим разностям $(m/\mu - m_1/\mu_1)$ и $(n/\mu - n_1/\mu_1)$, равным друг другу по величине и противоположным по знаку.

Как в формуле (84.2), я назову через p величину, мало отличную от m/μ, так что v окажется положительной или отрицательной переменной, очень малой относительно $\sqrt{\mu}$. Пусть также

$$p_1 = m_1/\mu_1 - (v_1/\mu_1)\sqrt{2m_1n_1/\mu_1}$$

будет лишь немного отлично от p, причём переменная v_1, положительная или отрицательная, очень мала относительно $\sqrt{\mu_1}$ и

$$(m_1/\mu_1 - m/\mu) = \delta,$$

где δ – малая дробь, также положительная или отрицательная. Тогда

$$p_1 - p = z = \delta + \frac{v}{\mu}\sqrt{\frac{2mn}{\mu}} - \frac{v_1}{\mu_1}\sqrt{\frac{2m_1n_1}{\mu_1}},$$

$$v_1 = (\delta - z)\mu_1\sqrt{\frac{\mu_1}{2m_1n_1}} + \frac{v\mu_1}{\mu}\sqrt{\frac{\mu_1 mn}{\mu m_1 n_1}}\,.$$

Если ε малая положительная дробь, и требуется определить вероятность того, что $p_1 \geq p + \varepsilon$, то переменной z следует придавать положительные значения, не меньшие ε. Бесконечно малые вероятности предыдущих значений p и p_1 будут равны Vdv и V_1dv_1, где V определяется формулой (84.1), а V_1 – той же формулой с заменой μ, m, n и v на μ_1, m_1, n_1 и v_1. Вероятность совместного появления этих значений будет равна $VdvV_1dv_1$, а требуемая вероятность окажется равной

$$\lambda = \int\int VV_1 dv dv_1.$$

Для упрощения я пренебрегаю вторым слагаемым формулы (84.1), и тогда

$$\lambda = \frac{1}{\pi}\int\int \exp(-v^2 - v_1^2) dv dv_1.$$

Если при этом интегрировании желательно заменить v_1 переменной z, то dv_1/dz следует вычислить, исходя из предыдущего значения v_1. Переменная v_1, как предполагается здесь, возрастает, и чтобы z также возрастало, необходимо изменить знак dv_1, так что

$$dv_1 = (\mu_1\sqrt{\mu_1}/\sqrt{2m_1n_1})dz.$$

[После длительных преобразований автор получил]

$$\lambda = \frac{1}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt,\ \lambda = 1 - \frac{1}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt \qquad (88.1a, b)$$

при $\varepsilon - \delta > 0$ и < 0. Здесь

$$\frac{(\varepsilon-\delta)\mu\mu_1\sqrt{\mu\mu_1}}{\sqrt{2(\mu^3 m_1 n_1 + \mu_1^3 mn)}} = \pm\, u, \qquad (88.2)$$

где u – положительная величина со знаками, соответствующими $\varepsilon - \delta > 0$ и < 0.

Нужно заметить, что при пренебрежении вторым слагаемым в формуле (84.1) пренебрегается и вероятность случая, при котором разность $p_1 - p$ точно равна ε, и λ оказывается вероятностью $p_1 - p > \varepsilon$, а не $\geq \varepsilon$. При $\varepsilon = \delta$ величина $u = 0$, и оба значения λ становятся равными 1/2.

Формулы (88.2) служат также для вычисления вероятности того, что неизвестный шанс p_1 превзойдёт заданную дробь. Действительно, я принимаю в [преобразованном] уравнении (88.2)

$\mu = \infty$, $m/\mu = w$, $\delta = (m_1/\mu_1) - w$.

Далее, для упрощения w заменяет $\varepsilon + w$, и μ, m, n будут подставлены вместо μ_1, m_1, n_1. Тогда

$$u = \pm\ (w - \frac{m}{\mu})\frac{\mu\sqrt{\mu}}{\sqrt{2mn}}. \qquad (88.3)$$

В соответствии с положительной или отрицательной разностью $w - m/\mu$ формулы (88.1a, b) выражают вероятность, что неизвестный шанс события, происшедшего m раз в очень большом числе $\mu = m + n$ испытаний, превзойдёт заданную дробь w.

89. Обращаясь к численным приложениям различных формул, я приведу для примера эксперимент Бюффона (§ 50). Событиями E и F будут появления орла и решётки при длительном подбрасывании монеты. По Бюффону, $m = 2048$, $n = 1992$, $\mu = 4040$ (m и n – количества наступлений этих событий при μ бросках). Подставляя эти значения в формулу (83.1) и полагая $u = 2$, мы получим

$$\frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt = 0.00468,\ R = 0.99555.$$

В то же время 0.50693 ± 0.02225 будут границами p, к которым относится указанная формула. Таким образом, неизвестный шанс p появления орла будет заключён между 0.48468 и 0.52918 с вероятностью 0.99555. Если желательно определить вероятность, что этот шанс превзойдёт 1/2, следует подставить предыдущие значения μ, m и n в формулу (88.3) и принять $w = 1/2$. И, выбрав знак *минус*, а потому и формулу (88.1b), мы получим

$u = 0.62298$, $\lambda = 0.81043$, $1 - \lambda = 0.18957$.

Это доказывает, что можно ставить несколько меньше пяти против одного за то, что шанс орла превосходит 1/2.

Эксперимент Бюффона можно разделить на две части с количествами испытаний 2048 и 1992. В первой из них орёл появился 1061 раз, а решётка – 987 раз; во второй части, соответственно, 987 и 1005 раз. Если же исходить из всех испытаний и применить формулу (87.2), можно также вычислить вероятность, что числа появлений орла и решётки должны содержаться в границах, указанных каждой частью эксперимента. Чтобы осуществить это, примем в указанной формуле и в соответствующих границах

$$m'/\mu' = m/\mu = 0.50693,\ n'/\mu' = n/\mu = 0.49307.$$

Другими словами, заменим m'/μ' и n'/μ', которые не предполагались известными, их приближёнными значениями, соответствующими всему эксперименту. Это возможно, если иметь в виду, что m' и n' входят лишь в члены порядка малости $1/\sqrt{\mu}$. Вместо μ следует также принять его полное значение, 4040.

Для первой части эксперимента $\mu' = 2048$; при $u = 2$ (см. выше) $w = 0.99558$ окажется вероятностью, что n', т. е. число появлений решётки, будет заключено в границы 1001 ± 79. Это условие действительно было выполнено. Для второй части $\mu' = 1992$ и w, снова при $u = 2$, равнялось 0.99560, т. е. вероятности, что число n' появлений решётки заключено в границы 982 ± 77, что также было выполнено. […]

Допустим, что мы не знаем, подбрасывалась ли одна и та же монета в обеих частях эксперимента, и по их результатам требуется установить, превосходила ли вероятность λ шанса орла в первой части тот же шанс во второй части на заданную дробь. Прежде всего, в уравнение (88.2) следует подставить

$$\mu = 1992,\ m = 987,\ n = 1005;$$
$$\mu_1 = 2048,\ m_1 = 1061,\ n_1 = 987.$$

Кроме того, $\delta = m_1/\mu_1 - m/\mu = 0.02257$, и это уравнение принимает вид

$$u = \pm\, 44.956\,(\varepsilon - 0.02257).$$

Если, к примеру, $\varepsilon = 0.02$, следует принять нижний знак и применить формулу (88.1b), так что

$$u = 0.11553,\ \lambda = 0.56589,\ 1 - \lambda = 0.43411$$

и можно ставить лишь 4 против трёх за то, что шанс орла в первой части превосходит его шанс во второй части на 1/50. При $\varepsilon = 0.025$ следует принять верхний знак и применить формулу (88.1a):

$u = 0.10925$, $\lambda = 0.43861$, $1 - \lambda = 0.56139$

и тогда можно ставить менее одного против одного за то, что эта разность более 1/40.

90. Я решаю здесь задачу, допускающую интересное приложение. Оно будет основано на предыдущих формулах и на лемме, которую я прежде всего изложу[8]. Урна содержит $c = a + b$ шариков, a белых и b чёрных. Из неё случайно извлекают l шариков без возврата, затем так же само $\mu = m + n$ других шариков. Я говорю, что при этом втором тираже вероятность появления m белых шариков и n чёрных не зависит ни от количества, ни от цвета ранее вышедших шариков[9] и равна вероятности их появления при $l = 0$.

Пусть $l + \mu$ тиражей произведено последовательно; обозначим через i общее число комбинаций $l + \mu$ шариков, которые могут быть извлечены; i' – число комбинаций μ последних шариков, m из которых белые и n – чёрные; и i_1– число комбинаций μ первых шариков, m из которых белые и n – чёрные. Шанс извлечь m белых и n чёрных после первого тиража равен i'/i, а шанс такого же результата до всех извлечений был равен i_1/i. Числа i' и i_1 равны друг другу, потому что вообще комбинации из l, а затем из μ определённых шариков, совпадают с комбинациями при обратном порядке тиражей. Поэтому равны также и дроби i'/i и i_1/i, а потому и вероятности, которые они представляют, ч. т. д.

Этот результат можно проверить следующим образом. Поскольку в урне содержалось a белых шариков и b чёрных, шанс появления m и n таких шариков при $m + n$ извлечениях есть функция $f(a, b, m, n)$. Аналогично, шанс появления g и h шариков в $g + h$ извлечениях будет тогда равен $f(a, b, g, h)$. Количество шариков в урне уменьшится до $a - g$ и $b - h$, и шанс появления m белых шариков и n чёрных в $\mu = m + n$ новых извлечениях будет равен $f(a - g, b - h, m, n)$. Произведение двух последних значений этой функции окажется равным шансу извлечь m белых шариков и n чёрных после появления указанных g белых и h чёрных[10].

Следовательно, суммируя $l + 1$ значений этого произведения, которые соответствуют всем целым и нулевым значениям g и h числом l, мы получим полное выражение шанса извлечь m белых и n чёрных шариков после появления каких-либо l шариков. Требуется установить, что этот шанс не зависит от l и равен $f(a, b, m, n)$, т. е. доказать, что

$$f(a, b, m, n) = \sum f(a, b, g, h)\, f(a - g, b - h, m, n),$$

где сумма распространяется от $g = 0$ и $h = l$ до $g = l$ и $h = 0$.

Я замечаю, что в соответствии с § 18

$$f(a,b,m,n) = \frac{\varphi(m,n)\ \varphi(a-m,b-n)}{\varphi(a,b)},\ \varphi(a,b) = C^{a}_{a+b},$$

$$f(a,b,g,h)f(a-g,b-h,m,n) = [\dots] =$$
$$\frac{\varphi(m,n)}{\varphi(a,b)}\varphi(g,h)\varphi(a-g-m,b-h-n).$$

Таким образом, после сокращения обеих частей этого уравнения на $\varphi(m, n)/\varphi(a, b)$ и учёта значения $f(a, b, m, n)$ проверяемое уравнение приобретает вид

$$\varphi(a-m,\ b-n) = \sum \varphi(g,h)\cdot\varphi(a-g-m,\ b-h-n),$$

а поскольку a и b произвольны, можно, если угодно, заменить их числами $a + m$ и $b + n$. Мы получим

$$\varphi(a,\ b) = \sum \varphi(g,h)\cdot\varphi(a-g,\ b-h).$$

Левая часть есть коэффициент при $x^a y^b$ в разложении $(x + y)^c$, правая же часть, снова при $x^a y^b$, является коэффициентом произведения разложений $(x + y)^l$ и $(x + y)^{c-l}$, т. е. при том же разложении, что и в левой части. Части уравнения тождественны, что и требовалось проверить.

91. Пусть числа $a, b, a - m, a - n$ очень велики. Приближенные значения $\varphi\ (m, n)$, $\varphi\ (a - m, b - n)$, $\varphi\ (a, b)$, а также $f(a, b, m, n)$ вычисляются при помощи формулы Стирлинга, и если ограничить эту формулу её первым членом и вычислить $f(a, b, m, n)$, то эту вероятность можно будет представить [после очень долгих преобразований] в виде

$$f(a,b,m,n) = H\exp(-t^2)[1 - \frac{4t^3(a-b)(c-2\mu)}{3\sqrt{2(c-\mu)\mu abc}}], \quad (91.1)$$

$$H = \sqrt{\frac{ab\mu(c-\mu)}{2\pi cmn(a-m)(b-n)}}$$

как шанс извлечь m белых шариков и n чёрных,

$$m = \frac{\mu a}{c} - \frac{t\sqrt{2(c-\mu)\mu abc}}{c^2}, \qquad (91.2a)$$

$$n = \frac{\mu b}{c} + \frac{t\sqrt{2(c-\mu)\mu abc}}{c^2}. \qquad (91.2b)$$

Разность $n - m$ окажется чётной при чётном μ и нечётной в противном случае. Пусть i – целое положительное число и $n - m = 2i$ или $2i - 1$, тогда величина t станет равной

$$t = 2i\delta + \gamma,\ \delta = \frac{c^2}{2\sqrt{2(c-\mu)\mu abc}},$$

$$\gamma = \frac{(a-b)\mu c}{2\sqrt{2(c-\mu)\mu abc}},\ \gamma = \frac{(a-b)\mu c}{2\sqrt{2(c-\mu)\mu abc}} - \delta$$

соответственно при чётном и нечётном μ. В результате подстановки указанного значения t в формулу (91.1) она выразит вероятность, что после μ последовательных тиражей число извлечённых чёрных шариков превзойдёт число белых на $2i$ или $2i - 1$. Принимая $i = 1, 2, 3, \ldots$ до тех пор, пока $\exp(-t^2)$ не станет неощутимой, или, если угодно, до $i = \infty$, сумма результатов окажется вероятностью, что после этих μ тиражей количество чёрных шариков превзойдёт количество белых на некоторое чётное или нечётное число единиц. Обозначая эту вероятность через s, мы получим

$$s = \sum H \exp(-t^2)[1 - \frac{4t^3(a-b)(c-2\mu)}{3\sqrt{2(c-\mu)\mu abc}}],$$

где сумма распространяется на все значения t от $\gamma + 2\delta$ до ∞, возрастающие по приращениям, равным 2δ. В соответствии с гипотезой δ – очень малая дробь, и сумма может выразить быстро сходящийся ряд, расположенный по степеням подобного приращения.

[После преобразований оказывается, что]

$$s = 1 - \frac{1}{\sqrt{\pi}} \int_v^\infty \exp(-t^2)dt - \Gamma \exp(-\gamma^2), \qquad (91.3a)$$

$$s = \frac{1}{\sqrt{\pi}} \int_v^\infty \exp(-t^2)dt - \Gamma \exp(-\gamma^2) \qquad (91.3b)$$

соответственно при $\gamma < 0$ и > 0 и положительном v, по величине равном γ, и при

$$\Gamma = \frac{(a-b)(c-2\mu)(7+4\gamma^2)+3c^2}{6\sqrt{2\pi(c-\mu)\mu abc}}.$$

Вероятность, что после μ тиражей $m = n = \mu/2$, что возможно лишь при чётном μ, равна

$$\sigma = \frac{c^2 \exp(-\gamma^2)}{\sqrt{2\pi(c-\mu)\mu abc}}. \qquad (91.4)$$

92. Пусть после μ тиражей было извлечено ещё μ', затем μ'', … других шариков, пока в урне не останется ничего. Тогда

$c = \mu + \mu' + \mu'' + \ldots$

Предположим также, что μ', μ'', … очень большие числа, равно как и μ, и обозначим через s', s'', … новые значения s, полученные после подстановки μ', μ'', … вместо μ по формулам (91.3a, b) в зависимости от того, было ли в начале тиражей число чёрных шариков b в урне больше или меньше, чем число белых, a, и, следовательно, стало ли $\gamma < 0$ или > 0. По лемме § 90 шансы извлечь больше чёрных шариков, чем белых, при μ, μ', μ'', … последовательных тиражах равны s, s', s'', … Эти шансы изменяются только ввиду неравенства μ, μ', μ'', …; они были бы тождественны, будь одинаковы эти последние. Обозначим среднее значение s, s', s'', … через r[11]. Пусть общее число всех тиражей α очень велико и в j из них количество чёрных шариков превосходило количество белых. Тогда вероятность величине j находиться внутри заданных границ будет, ввиду первого предложения § 52, той же, что и при равенстве всех шансов s, s', s'', … друг другу и их среднему r. Поэтому, подставив в формулу (79.1) α, r, $1 - r$ вместо μ, q, p, мы получим вероятность

$$R = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt + \frac{1}{\sqrt{2\pi\alpha r(1-r)}}\exp(-u^2)$$

того, что j будет находиться в границах $\alpha r \mp u\sqrt{2\alpha r(1-r)}$ или окажется равной какой-либо из них. Здесь *u* мало по сравнению с $\sqrt{\alpha}$.

Таково решение предложенной нами задачи. Её можно применить к выборам депутатов большой страны, например Франции. Вот в чём она состоит. Число избирателей во всей Франции обозначим через *c*; среди них *a* придерживаются одного мнения, а $b = c - a$ – противоположного мнения. Распределим их по α избирательным округам, в каждом из которых по большинству голосов избирается один депутат. Требуется определить вероятность *R*, того, что число j депутатов, придерживающихся второго мнения, будет заключено в заданные границы, если в округах состоят $\mu, \mu', \mu'', \ldots$ избирателей.

Пусть границами j будут только что указанные, тогда требуемая вероятность *R* выразится предыдущей формулой. Каждый округ состоит из избирателей, проживающих в одном и том же районе, а не выбранных случайно из их общего списка, как мы предположили. И всё же будет полезно узнать, что произойдёт в соответствии с этой гипотезой, что мы и покажем на примерах.

93. Во Франции число избирательных округов, равное числу депутатов, составляет 459, а общее число избирателей можно оценить как примерно равное 200 000[12] Я предположу, что числа $\mu, \mu', \mu'', \ldots$ совпадают и что μ нечётно:

$$\alpha = 459,\ \mu = 435,\ c = \alpha\mu = 199\,665.$$

Пусть также $a = 94\,835$ и $b = 104\,830$, т. е. что разность этих чисел составляет почти двадцатую часть числа избирателей. Величина γ будет отрицательна и $v = -\gamma$. Принимая второе из значений γ из § 91, получим

$$v = 0.77396,\quad \frac{1}{\sqrt{\pi}}\int_v^\infty \exp(-t^2)dt = 0.13684,$$

а ввиду формулы (91.3а) $s = 0.85426$, $1 - s = 0.14574$.

Шанс избрания депутата избирателями более многочисленной партии поэтому превысит 21/25, а меньшинство, хоть и ненамного отличающееся от большинства, никак не может надеяться избрать более 4/25 депутатов. Подставляя значения s и $1 - s$ вместо r и $1 - r$ в выражение для R из § 92 при $\alpha = 459$ и $u = 2$, мы найдём, что $R = 0.99682$ для вероятности, что число депутатов, избранных более многочисленной партией, окажется в границах 392 ∓ 21, и в границах 67 ± 21 для другой партии. Протяжённость границ значительна, потому что это α не исключительно велико.

Я неизменно предполагаю, что разность $b - a$ составляет почти 1/20 от c. Но пусть μ чётно, тогда

$$\alpha = 459,\ \mu = 436,\ c = \alpha\mu = 200\ 124.$$

Пусть также $a = 95\ 064$, $b = 105\ 060$. По-прежнему $v = -\gamma$, но для γ следует принять его первое значение в § 91, так что

$$v = 0.74006,\ \frac{1}{\sqrt{\pi}}\int_v^\infty \exp(-t^2)dt = 0.14764,$$

и $s = 0.84279$, $1 - s = 0.15721$.

Поскольку μ чётно, возможен случай $m = n$. По формуле (91.4) его шанс равен $\sigma = 0.022178$. Добавив $\sigma/2$ к s, мы получим $s = 0.85388$, что лишь очень ненамного меньше того, который соответствует нечётному μ.

Чтобы указать на влияние неравенства количества избирателей в округах, я предположу, что половина их них поровну распределена по 1/3 округам, а вторая половина – по оставшимся округам. Для первой трети округов будет

$$\alpha/3 = 153,\ \mu = 654,\ \alpha\mu/3 = 100\ 062,$$

а для остальных третей

$$2\alpha/3 = 306,\ \mu = 327,\ 2\alpha\mu/3 = 100\ 062.$$

Кроме того, я предположу, что $a = 95\ 062$, $b = 105\ 062$, $c = 200\ 124$, так что неравенство численностей партий составит, как и прежде, почти 1/20 числа избирателей. В первом случае μ чётно, во втором случае – нечётно и соответственно

$s = 0.89429$, $\sigma = 0.01376$, $s + \sigma/2 = 0.90117$; $s = 0.81981$,

так что средний шанс избрания депутата от партии большинства

$r = (0.90117 + 0.81981)/2 = 0.86049$.

Он немного превосходит тот, который имел место при одном и том же числе избирателей во всех округах. Но если разность $b - a$ начнёт возрастать, шанс избрания депутата от партии меньшинства начнёт очень быстро уменьшаться и скоро почти исчезнет. Чтобы показать это, я предположу, что избиратели поровну распределены по округам и приму те же значения α, μ и c, что и в первом примере. Кроме того, пусть $a = 89\,855$, $b = 109\,830$, так что разность $b - a$ станет равной почти $c/10$, т. е. вдвое большей, чем в том примере. Тогда

$s = 0.98176$, $1 - s = 0.01824$

и шанс избрания депутата от партии меньшинства оказывается не большим, чем почти 1/60[13]. Ввиду малости s следует определять вероятность P того, что при подобных выборах количество депутатов от партии меньшинства не превзойдёт заданного числа n, по формуле § 81. И оказывается, что

$w = \alpha(1 - s) = 8.3713$, $n = 15$, $P = 0.98713$, $1 - P = 0.01287$.

Если же увеличить разность численностей партий до 30 000, т. е. до 3/20 общего числа избирателей, то шанс $1 - s$ станет меньше 1/1000, так что весьма вероятно, что от партии меньшинства не будет тогда избран ни один депутат.

Если это так, то представительный строй оказывается лишь обманом, потому что меньшинство, составляющее 90 000 избирателей из 200 000, будет представлено лишь очень небольшим числом депутатов, а меньшинство в 85 000 будет иметь лишь очень слабый шанс одного-единственного выразителя своего интереса в палате депутатов. И достаточно 3/20 избирателей в период между двумя сессиями изменить своё мнение на противоположное, чтобы вся палата последовала их примеру.

Но избиратели, из которых состоит каждый округ, не отобраны случайно, как мы это предположили, из их общего списка по всей

Франции. В каждом районе преобладающее мнение формируется и поддерживается частными причинами, как, например, местными интересами, влиянием правительства и некоторых граждан. Всё же полезно указать на исключительную изменчивость, которую случай может произвести в составе палаты депутатов очень небольшими изменениями в соотношении избирателей, придерживающихся противоположных мнений.

Примечания

1. Вместо обозначения $n!$, которого у Пуассона ещё не было, он иногда вводил два других, одно из которых мы ниже сохранили. В этом параграфе непосредственно нужна лишь формула Стирлинга (67.3), но на некоторые промежуточные преобразования Пуассон ссылался в дальнейшем.

2. Ряд Стирлинга расходится (Фихтенгольц 1947/1951, с. 820).

3. Мы бы сказали иначе: в прямой, но не в обратной задаче p и q известны. Это поясняет, почему дисперсии соответствующих случайных величин различны. В обратной задаче она больше, что стало ясно Бейесу, который еще не владел понятием дисперсии (Шейнин 2007) и о котором весьма положительно отозвался Лаплас (1814/1999, с. 862, левый столбец), хотя и не раскрыл суть его работы. Дисперсию ввёл Гаусс в 1823 г., но Пуассон почти не применял её. В одном случае он (1837, с. 73) фактически воспользовался ей для оценки качества огнестрельного оружия.

4. Приведенная формула встречалась у Монмора (Montmort 1708/1713, p. 245), Todhunter (1865/1965, p. 97).

5. Заметим, что $\int_{-\infty}^{t}(x-a)^r\varphi(x)dx$, где $\varphi(x)$ – соответствующая плотность распределения, называется неполным моментом r-го порядка. Ниже Пуассон вычислил интеграл от $t^{2i}\exp(-t^2)$ в пределах 0 и ∞ при различных значениях i, Гаусс же (1816) установил обычные, *полные* моменты нормального распределения. Соотношение § 82 книги Пуассона и этой работы Гаусса заслуживает изучения, но сразу укажем, что упорное нежелание Пуассона воспользоваться результатами великого учёного обернулось против него.

6. В формуле (85.2) нет $\alpha\sqrt{\mu'}$.

7. В указанном месте было принято u, равное четырём или пяти.

8. После публикации моего примечания [см. наше Прим. 18 к гл. 1. О. Ш.] мне заметили, что содержащееся в нём предложение имелось в лемме, которую я уже применил при решении задачи об игре в тридцать и сорок (1825, с. 70). Автор. Указанная страница неверна. О. Ш.

9. Пуассон таким образом рассматривал субъективные вероятности. См. также конец Прим. 18 к гл. 1.

10. Пояснения недостаточны. Чуть ниже указано, что $l = g + h$.

11. Можно заключить, что $\alpha = c$. Изложение, стало быть, не вполне продумано.

12. Избиратели, стало быть, составляли лишь небольшую долю населения.

13. Точнее, 1/55.

Библиография

Курно О. (1843, франц.), *Основы теории шансов и вероятностей*. М., 1970.
Фихтенгольц Г. М. (1947), *Курс дифференциального и интегрального исчисления*, т. 2. М.-Л., 1951. Второй том трёхтомника.
Шейнин О. Б. (2007), К истории теоремы Бейеса. *Историко-математич. исследования*, вып. 12 (47), с. 312 – 320.
Gauss C. F., Гаусс К. Ф. (1816 нем.), Определение точности наблюдений. В книге автора *Избр. геод. соч*., т. 1. М., 1957, с. 121 – 128.
Kramp Chr. (1798), *Analyse des réfractions astronomiques*. Leipzig – Strassburg. Курно (1843) перепечатал его таблицу нормального распределения.
Laplace P. S., Лаплас П. С. (1814 франц.), *Опыт философии теории вероятностей*. В книге Прохоров Ю. В., редактор (1999), *Вероятность и математическая статистика. Энциклопедия*. М., с. 834 – 866.
Montmort P. R. (1708), *Essai d'analyse sur les jeux de hasard*. Paris, 1713; New York, 1980.
Poisson S. D. (1825), Sur l'avantage du banquier au jeu de trente-et-quarante. *Ann. Math. Pures et Appl.*, t. 16, pp. 173 – 208.
--- (1830), Sur la proportion des naissances des filles et des garçons. *Mém. Acad. Sci. Paris*, t. 9, pp. 239 – 308.
--- (1837), Sur la probabilité du tir a la cible. *Mémorial d'artill.*, No. 4, pp. 59 – 94.
Todhunter I. (1865), *History of the Mathematical Theory of Probability*. New York, 1949, 1965.

Глава 4. Продолжение исчисления вероятностей, зависящих от очень больших чисел

Опечатки/ошибки, не замеченные автором

1. В § 94, с. 247 оригинала, в первом из трёх, содержащихся в этом параграфе интегралов, отсутствует дифференциал dx.

2. В § 95, с. 251 оригинала, ни в одном интеграле в первой выключенной строке нет пределов, а в первом из них нет дифференциала (в переводе вставлен).

3. В § 96, с. 253 оригинала, в третьей выключенной формуле подынтегральная функция должна быть $\exp(-t^2)$.

4. В § 103, с. 274 оригинала. В первом из интегральных выражений для γ_i результат интегрирования ошибочен: вместо γ_i указано γ.

5. В § 104, с. 277 оригинала, в строке 10 шанс А следует заменить на шанс Е.

6. В § 106, с. 285 оригинала, после второй выключенной формулы, упоминается $1/\mu$ в знаменателе вместо μ в знаменателе.

7. В § 106, с. 286 оригинала, автор ввёл величину λ и включил её в три равенства. В левой части второго из них вставлена не имеющая смысла разность $s/\mu - s^2/\mu^2$.

8. В § 106, с. 287 оригинала, в третьей строке после первой выключенной формулы, порядок пределов интеграла указан неверно.

9. В § 107, с. 288 оригинала. Во втором уравнении второй выключенной формулы в знаменателе левой части должно быть μ', а не μ.

10. В § 109, с. 295 оригинала, в строке 4 снизу порядок малости δ должен быть $1/\sqrt{\mu}$, а не μ, см. строку 4 снизу с. 296.

11. В § 110, с. 301 оригинала, строка 5 снизу. Наибольшее целое число, содержащееся в s следует заменить на содержащееся в α.

12. В § 112/2, с. 308 оригинала, строка 2 после формулы (b). Шансы F и F следует заменить на шансы E и F.

13. В § 112/8, с. 312 оригинала, в первой скобке формулы для Г явно должно быть слагаемое $-4g$. В той же формуле у последней скобки отсутствует показатель степени.

14. В § 112/12, с. 315 оригинала, строка 2 после первой выключенной формулы. Величину γ следует заменить на величину А.

15. В начале § 113, с. 316 оригинала. Исходные уравнения Лапласа упомянуты лишь частично и притом ошибочно. См. Прим. 18.

Все возможные исправления внесены в перевод.

94. Мы теперь займёмся формулами, относящимися к переменным шансам. Это приведёт нас к доказательству трёх основных предложений, указанных в §§ 52 и 53, и имевших следствием *закон больших чисел*.

Рассмотрим ряд последовательных $\mu = m + n$ испытаний, в течение которых шансы противоположных событий E и F каким-то образом менялись, будучи равными p_1 и q_1 в первом испытании, p_2 и q_2 – во втором испытании, …, p_μ и q_μ – в последнем. Имеют место равенства

$$p_1 + q_1 = p_2 + q_2 = \ldots = p_\mu + q_\mu = 1.$$

Обозначим через U вероятность, что указанные события наступят в каком-то порядке m и n раз. По правилу § 20 U является коэффициентом при $u^m v^n$ в разложении произведения[1]

$$X = (up_1 + vq_1)\,(up_2 + vq_2)\,\ldots\,(up_\mu + vq_\mu). \qquad (94.1)$$

Если $u = e^{x\sqrt{-1}}$, $v = e^{-x\sqrt{-1}}$, то слагаемое $Uu^m v^n$ этого произведения запишется в виде $Ue^{(m-n)x\sqrt{-1}}$. Все остальные слагаемые будут включать экспоненты, отличающиеся от указанной, так что при умножении произведения на $e^{-(m-n)x\sqrt{-1}}$ и интегрировании в пределах – π и π все остальные сомножители исчезнут:

$$\int_{-\pi}^{\pi} X \exp[-(m-n)x\sqrt{-1}]dx = 2\pi U.$$

[…] Сомножители в (94.1) равны

$$up_i + vq_i = \cos x + (p_i - q_i)\sin x\sqrt{-1}.$$

Примем

$$\cos^2 x + (p_i - q_i)^2 \sin^2 x = \rho_i^2$$

и введём действительный угол r_i

$$(1/\rho_i)\cos x = \cos r_i,\ (1/\rho_i)(p_i - q_i)\sin x = \sin r_i,$$

так что

$$up_i + vq_i = \rho_i \exp(r_i\sqrt{-1}).$$

Величина ρ_i двузначна; будем считать её положительной и примем для сокращения письма

$$\rho_1\rho_2\ldots\rho_\mu = Y,\ r_1 + r_2 + \ldots + r_\mu = y. \qquad (94.2a, b)$$

Тогда *u* примет вид […]

$$U = \frac{1}{2\pi}\int_{-\pi}^{\pi} Y\cos[y-(m-n)x]dx + \frac{\sqrt{-1}}{2\pi}\int_{-\pi}^{\pi} Y\sin[y-(m-n)x]dx.$$

Эта сумма сводится к своему первому слагаемому, и результат обозначен как формула (94.3). Интегрирование по обычным правилам возможно в конечном виде. Но если μ не будет большим числом, приведенная формула окажется бесполезной для подсчёта значения U. Когда же это число очень велико, можно, как будет видно, вывести из этой формулы сколь угодно приближённое значение U.

95. При $x = 0$ каждый сомножитель в формуле (94.2a) оказывается равным единице, и ещё меньшим при всех других значениях x внутри пределов интегрирования. Стало быть, когда μ очень велико, произведение в формуле (94.2a), вообще говоря, очень мало при всех, кроме очень небольших значений x, так что, если μ становится бесконечным, Y исчезает при всех конечных x.

Есть одно исключение, при котором все сомножители Y безгранично стремятся к единице; как известно, произведение бесконечного числа таких сомножителей может оказаться конечным. Ввиду того, что

$$\rho_i^2 = 1 - 4p_iq_i\sin^2 x,$$

один из шансов событий E и F или их произведения p_iq_i неограниченно убывает во время испытаний. Исключая упомянутый частный случай, можно, если μ очень большое число, считать переменную x очень малой величиной и пренебречь в предыдущем интеграле той его частью, которая соответствует другим значениям x.

Поэтому следующие ряды, расположенные по степеням x^2, быстро сходятся:

$$\rho_i = 1 - 2p_iq_ix^2 + [(2/3)p_iq_i - 2p_i^2q_i^2)x^4 - \dots],$$
$$\ln\rho_i = - 2p_iq_ix^2 + [(2/3)p_iq_i - 4p_i^2q_i^2]x^4 - \dots,$$
$$\ln Y = -\mu k^2x^2 + \mu[(1/3)k^2 - k'^2]x^4 - \dots$$

Здесь было принято

$$\mu k^2 = 2\sum p_iq_i,\ \mu k'^2 = 4\sum p_i^2q_i^2,\ \dots$$

и суммы распространялись от $i = 1$ до $i = \mu$.

Полагая также, что $x = z/\sqrt{\mu}$, рассматривая новую переменную z как величину, очень малую сравнительно с $\sqrt{\mu}$, и пренебрегая величинами порядка малости $1/\mu$, мы получаем

$$Y = \exp(- k^2z^2).$$

Кроме того, в соответствии со значениями ρ_i и $\sin r_i$,

$$r_i = (p_i - q_i)x + (4/3)(p_i - q_i)p_iq_ix^3 + \dots$$

Введём обозначение

$$h = (4/3\mu)\sum(p_i - q_i)p_iq_i$$

и средние значения p и q шансов событий E и F, так что $p + q = 1$.

Если сохранять только величины порядка малости $1/\sqrt{\mu}$, то

$$y = z(p - q)\sqrt{\mu} + z^3h/\sqrt{\mu},$$
$$\cos[y - (m - n)x] = \cos(zg\sqrt{\mu}) - [z^3h/\sqrt{\mu}]\sin(zg\sqrt{\mu}),$$

где

$$g = (p - m/\mu) - (q - n/\mu).$$

Я подставляю эти значения Y и $\cos[y - (m - n)x]$ в окончательную формулу для U (94.3), принимая при этом, что $dx = dz/\sqrt{\mu}$:

$$U = \frac{2}{\pi\sqrt{\mu}}[\int \exp(-k^2z^2)\cos(zg\sqrt{\mu})dz \; -$$

$$\frac{h}{\sqrt{\mu}}\int \exp(-k^2z^2)z^3\sin(zg\sqrt{\mu})dz].$$

[После преобразований оказывается, что] вероятность

$$U = \frac{1}{k\sqrt{\pi\mu}}\exp(-\theta^2) - \frac{h\theta\exp(-\theta^2)}{2k^4\mu\sqrt{\pi}}(3+2\theta^2), \qquad (95.1)$$

где

$p - m/\mu = \theta k\sqrt{\mu}$, $q - n/\mu = -\theta k\sqrt{\mu}$, так что $g = 2\theta k\sqrt{\mu}$,

соответствует равенствам

$m = p\mu - \theta k\sqrt{\mu}$, $n = q\mu + \theta k\sqrt{\mu}$,

т. е. значениям, почти пропорциональным средним шансам p и q и числу μ испытаний[2].

96. Поскольку m и n целые числа, θ должно быть кратно $\delta = 1/k\sqrt{\mu}$ или равно нулю. При $\theta = 0$ формула (95.1) укажет вероятность $1/k\sqrt{\pi\mu}$ того, что в точности $m/n = p/q$. Обозначим через t положительную величину, кратную δ; последовательно положим в этой формуле $\theta = -t$ и t и сложим полученные результаты, получая сумму

$$\frac{2}{k\sqrt{\pi\mu}}\exp(-t^2), \qquad (96.1)$$

т. е. вероятность, что m и n будут равны соответственно одному из значений

$$p\mu \mp kt\sqrt{\mu},\; q\mu \pm kt\sqrt{\mu}.$$

Обозначим через u заданное кратное δ и в сумме (96.1) примем последовательно $t = \delta, 2\delta, \ldots, u$. Обозначим сумму полученных таким образом результатов, увеличенную на значение U при $\theta = 0$, через R:

$$R = \frac{1}{k\sqrt{\pi\mu}} + \frac{2}{k\sqrt{\pi\mu}}\sum \exp(-t^2).$$

Она равна вероятности, что числа m и n заключены в границах

$$p\mu \mp uk\sqrt{\mu},\ q\mu \pm uk\sqrt{\mu}$$

или равны какой-либо из них. Сумма распространяется на значения t от $t = \delta$ до u и возрастает по приращениям, равным δ. Но её можно заменить разностью сумм экспоненциальной функции от $t = \delta$ до ∞ и от $t = u + \delta$ до ∞. При помощи формулы Эйлера, которую мы уже применили в § 91, эта последняя сумма, умноженная на δ, с принятой нами степенью аппроксимации, т. е. с пренебрежением квадратом δ, оказывается равной

$$\int_u^\infty \exp(-t^2)dt - \frac{\delta}{2}\exp(-u^2).$$

При $u = 0$ сумма, распространенная от $t = \delta$ до ∞, умноженная на δ, будет равна $(1/2)(\sqrt{\pi} - \delta)$. Следовательно, вычитая эту последнюю величину из предыдущей, и разделив разность на δ, мы получим для суммы в выражении для R

$$\sum \exp(-t^2) = \frac{1}{2\delta}\sqrt{\pi} - \frac{1}{\delta}\int_u^\infty \exp(-t^2)dt - \frac{1}{2} + \frac{1}{2}\exp(-u^2).$$

С учётом значения δ это выражение принимает вид

$$R = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt + \frac{\exp(-u^2)}{k\sqrt{\pi\mu}}. \qquad (96.2)$$

Если шансы p_i и q_i постоянны и поэтому равны своим средним значениям p и q, то $k = \sqrt{2pq}$. Формула (96.2) и предыдущие границы для m и n совпадут соответственно с формулой (79.1) и теми границами. При необходимости указанное совпадение результатов, полученных столь различными методами, может послужить подтверждением наших вычислений.

Если принять для u незначительную величину, например, 3 или 4, значение R станет очень близким единице. И поэтому почти достоверно, что при очень большом числе μ испытаний отношения m/μ и n/μ будут очень немного отличаться от средних шансов p и q, станут ближе приближаться к ним по мере дальнейшего возрастания μ и, наконец, точно совпадут с ними, если μ сможет стать бесконечным. В этом состоит первое из двух общих предложений § 52.

97. Пусть теперь А – некоторая вещь, способная принимать многие положительные и отрицательные значения, которые мы будем считать кратными заданной величине w. Эти значения заключены в границы от αw до βw включительно, так что их число равно $\beta - \alpha + 1$. Здесь α и β – целые числа или нули, притом без учёта знака второе из них превосходит первое. Если же А может принимать лишь одно значение, то $\beta = \alpha$. При каждом испытании, произведенном для установления А, его возможные значения не равновероятны. Кроме того, для большей общности мы предположим, что шансы одного и того же значения изменяются от одного испытания к другому.

Пусть n будет каким-либо числом, заключённым между α и β или равным какой-либо из этих величин. Далее, обозначим через N_1 шанс значения nw величины А в первом испытании, через N_2 – этот шанс во втором испытании, и т. д. и пусть s будет суммой значений А, которые имеют место при μ последовательных испытаниях. Требуется определить вероятность, что эта сумма находится в заданных границах. Обозначим прежде всего вероятность точного равенства $s = mw$ через П. Здесь m – заданное число, *находящееся между α и β* или равное какой-либо из этих величин. [Автор исключил из текста выделенные нами слова, но после такого исправления фраза оказалась непонятной.]

Составим произведение

$$\sum N_1 t^{nw} \sum N_2 t^{nw} \ldots \sum N_\mu t^{nw},$$

в котором t – неопределённая величина, а суммы распространяются на все значения n от α до β. Если разложить это произведение по степеням t^w, то легко будет усмотреть, что П окажется в нём коэффициентом при t^{mw}. Это ясно для случая $\mu = 1$. Пусть теперь $\mu = 2$. Обозначим через $n'w$ и $n''w$ экспоненты t, которые эта величина примет в соответствующих суммах; ясно, что значение А, равное mw, может появиться столькими различными способами, сколько различных решений имеет уравнение $n'+n''=m$ при n' и n'', заключёнными между α и β. Вероятность каждого способа равна произведению значений N_1 и N_2, которые соответствуют каждой паре чисел n' и n''. Следовательно, полная вероятность равенства $s = mw$ выражается коэффициентом при t^{mw} в произведении этих двух первых сумм. Указанное соображение без труда распространяется на случаи $\mu = 3, 4, \ldots$ Если все величины $N_1, N_2, \ldots$ совпадают, их произведение становится степенью μ одного из многочленов, соответствующих суммам $\sum$, и этот случай был рассмотрен в § 17.

Подобным же рассуждением, полагая $t^w = \exp(\theta\sqrt{-1})$ и обозначая через X произведение μ сумм $\sum$, мы получим[3]

$$\Pi = \frac{1}{2\pi}\int_{-\pi}^{\pi} X\exp(-m\theta\sqrt{-1})d\theta.$$

Пусть i и i_1– два данных числа и P – вероятность того, что сумма s заключена между iw и i_1w или равна какой-либо из этих границ. Тогда значение P можно установить по Π, принимая последовательно $m = i, i + 1, \ldots, i_1$. И сумма значений, соответствующих $\exp(-m\theta\sqrt{-1})$, будет выражена следующим образом: […].

Для упрощения я предполагаю, что w бесконечно мало, а i и i_1 – бесконечные числа, и я приму, что

$$iw = c - \varepsilon,\ i_1w = c + \varepsilon,\ \theta = wx,\ d\theta = wdx,$$

где c и ε – заданные константы, притом ε положительно, чтобы соблюдалось неравенство $i_1 > i$, которое подразумевается в выражении P. Пределы интеграла относительно новой переменной x бесконечны. Далее, $\sin(\theta/2) = wx/2$, и, если пренебречь числами $\pm$ 1/2 сравнительно с i и i_1, P примет вид

$$P = \frac{1}{\pi}\int_{-\infty}^{\infty} X\exp(-cx\sqrt{-1})\sin\varepsilon x\frac{dx}{x}. \qquad (97.1)$$

Возможные значения А возрастают бесконечно малыми приращениями, и их число следует считать бесконечным, а вероятность каждого – бесконечно малой. Если обозначить данные константы через a и b, а непрерывную переменную – z, то

$$\alpha w = a,\ \beta w = b,\ nw = z,\ t^{nw} = \exp(xz\sqrt{-1}).$$

В то же время пусть

$$N_1 = wf_1z,\ N_2 = wf_2z,\ \ldots$$

Каждая из сумм, содержащаяся в X, становится определённым интегралом с пределами, равными a и b. Принимая w за dz, мы заключаем, что

$$X = \int_a^b e^{xz\sqrt{-1}} f_1zdz \int_a^b e^{xz\sqrt{-1}} f_2zdz \ldots \int_a^b e^{xz\sqrt{-1}} f_\mu zdz \qquad (97.2)$$

оказывается произведением μ сомножителей, которое следует подставить в формулу (97.1) вместо X.

98. Эта формула выражает вероятность, что после μ испытаний сумма значений А будет заключена между данными величинами $c - \varepsilon$ и $c + \varepsilon$. При n-м испытании бесконечно малый шанс значения z будет равен f_nzdz, притом все возможные значения А заключаются по предположению между границами a и b, и одно из них достоверно имеет место при каждом испытании. Поэтому

$$\int_a^b f_nzdz = 1.$$

Функция f_nz может быть непрерывной или разрывной, но в границах между a и b только положительной.

Если шанс каждого значения z в течение испытаний остаётся постоянным, эта функция будет независимой от n. Обозначив её через fz, мы будем иметь

$$X = \left(\int_a^b e^{xz\sqrt{-1}} fzdz\right)^\mu,\ \int_a^b fzdz = 1.$$

Если же все значения А равновероятны, fz окажется константой, и, поскольку она должна удовлетворять последнему уравнению, то будет равна $1/(a - b)$.

Пусть $a = h - g,\ b = h + g$, тогда

$$fz = \frac{1}{2g},\ \int_a^b e^{xz\sqrt{-1}} fzdz = \frac{\sin gx}{gx} e^{hx\sqrt{-1}},$$

и формула (97.1) примет вид […]

$$P = \frac{2}{\pi}\int_0^\infty \left(\frac{\sin gx}{gx}\right)^\mu \frac{\sin \varepsilon x}{x}\cos[(\mu h - c)x]dx.$$

[После длительных преобразований автор получил в § **99**]

$$2(2g)^\mu P = \frac{\Gamma - \Gamma_1}{\mu!}, \qquad (99.1)$$

$$\Gamma = \pm(\gamma + \mu g + \varepsilon)^\mu \mp \mu(\gamma + \mu g - 2g + \varepsilon)^\mu \pm$$
$$C_\mu^2(\gamma + \mu g - 4g + \varepsilon)^\mu \mp C_\mu^3(\gamma + \mu g - 6g + \varepsilon)^\mu + \ldots,$$

Γ_1 же равно Γ с изменённым знаком ε. Уравнение (99.1), как и требовалось, представляет значение P в конечной форме.

100. В случае $\mu = 1$, т. е. одного-единственного наблюдения, P является вероятностью того, что значение А, которое в соответствии с гипотезой должно содержаться внутри заданных границ a и b или $h - g$ и $h + g$, после наблюдения оказывается в границах $c - \varepsilon$ и $c + \varepsilon$. Если эти последние заключают в себе первые, то должно выполняться равенство $P = 1$. Если же первые заключают в себе последние, то P должно быть отношением интервала 2ε последних к интервалу $2g$ первых. Далее, если обе последние границы оказываются вне интервала первых, то обязательно $P = 0$; если $c - \varepsilon$ находится внутри интервала между $h - g$ и $h + g$, а $c + \varepsilon$ – вне его, то P должно быть равно отношению разности $(h + g) - (c - \varepsilon)$ к интервалу $2g$. Наконец, если $c + \varepsilon$ находится в интервале $h - g$, $h + g$, а $c - \varepsilon$ – вне его, то P должно быть равно отношению разности $(c + \varepsilon) - (h - g)$ к тому же интервалу. Вот эти различные значения P:

$$P = 1,\ P = \frac{\varepsilon}{g},\ P = 0,\ P = \frac{h + g - c + \varepsilon}{2g},\ P = \frac{c + \varepsilon - h + g}{2g}.$$

Они выводятся из уравнения (99.1), которое при $\mu = 1$ принимает вид

$$P = \frac{1}{4g}(\Gamma - \Gamma_1).$$

Кроме того, $\gamma = h - c$, так что

$$\Gamma = \pm(h+g-c+\varepsilon) \mp (h-g-c+\varepsilon), \qquad (100.1a)$$
$$\Gamma_1 = \pm(h+g-c-\varepsilon) \mp (h-g-c-\varepsilon). \qquad (100.1b)$$

В первом случае из пяти мы имеем $c + \varepsilon > h + g$ и $c - \varepsilon < h - g$. В формуле (100.1a) величины в скобках положительны, а в (100.1b) – отрицательны. Поэтому в первой из этих формул следует принять верхние знаки, и нижние знаки во второй формуле.

Во втором случае мы имеем $h + g > c + \varepsilon$, $h - g < c - \varepsilon$. Перед первыми слагаемыми обеих формул следует принять верхние знаки, а перед вторыми – нижние знаки.

В третьем случае $h - g > c + \varepsilon$ и в обеих формулах следует принять верхние знаки. Но здесь возможно и условие $h + g < c - \varepsilon$, и тогда следует принять нижние знаки и $P = 0$.

В четвёртом случае $c - \varepsilon > h - g$, $c - \varepsilon < h + g$, $c + \varepsilon > h + g$. Во второй формуле следует принять нижние знаки, а в первой формуле – верхний знак перед первым слагаемым и нижний знак – перед вторым.

В пятом случае $c - \varepsilon < h - g$, $c + \varepsilon > h - g$, $c + \varepsilon < h + g$. В первой формуле следует принять верхние знаки, также и перед первым слагаемым второй формулы, и нижний знак перед вторым слагаемым.

Вот результаты всех указанных случаев:

1. $\Gamma = 2g,\ \Gamma_1 = -2g,\ P = 1.$

2. $\Gamma = 2h - 2c + 2\varepsilon,\ \Gamma_1 = 2h - 2c - 2\varepsilon,\ P = \varepsilon/g.$

3. $\Gamma = 2g,\ \Gamma_1 = 2g,\ P = 0.$

4. $\Gamma = 2h - 2c + 2\varepsilon,\ \Gamma_1 = -2g,\ P = \dfrac{h+g-c+\varepsilon}{2g}.$

5. $\Gamma = 2g,\ \Gamma_1 = 2h - 2c - 2\varepsilon,\ P = \dfrac{c+\varepsilon-h+g}{2g}.$

Значения P при одном-единственном наблюдении можно также проверить по формуле (97.1) для общего случая. Если считать $f_1 z$ разрывной функцией, равной нулю при всех значениях z вне заданных границ a и b, то вероятность P того, что значение А должно содержаться в границах $c - \varepsilon$ и $c + \varepsilon$ будет очевидно равна

$$P = \int_{c-\varepsilon}^{c+\varepsilon} f_1 z dz .$$

При μ = 1 по формулам (97.2) и (97.1)

$$X = \int_a^b e^{xz\sqrt{-1}} f_1 z dz ,$$

и […]

$$P = \frac{1}{\pi}\int_a^b \left(\int_0^\infty \frac{\sin[(c+\varepsilon-z)x]}{x} dx - \int_0^\infty \frac{\sin[(c-\varepsilon-z)x]}{x} dx \right) f_1 z dz .$$

Но $\int_0^\infty \frac{\sin \gamma x}{x} dx = \pm \frac{\pi}{2}$

при положительной или отрицательной константе γ. Поэтому указанная разность интегралов равна нулю или π при одних и тех же или противоположных знаках $c + \varepsilon - z$ и $c - \varepsilon - z$ и интеграл по z будет равен нулю для всех значений z, либо больших, чем $c + \varepsilon$, либо меньших, чем $c - \varepsilon$. Он должен распространяться одновременно между a и b, и между $c - \varepsilon$ и $c + \varepsilon$. А так как мы полагаем, что $f_1 z$ равна нулю вне первых границ, то значение P сводится к интегралу от $f_1 z$ в пределах от $z = c - \varepsilon$ до $c + \varepsilon$, что и требовалось проверить.

101. Если μ – очень большое число, мы можем при помощи преобразований, подобных проведенным в § 95, заменить формулу (97.1) другой, которая установит приближённое значение P. Прежде всего заметим, что формулу (97.2) можно записать в виде

$$X = \int_a^b e^{xz_1\sqrt{-1}} f_1 z_1 dz_1 \int_a^b e^{xz_2\sqrt{-1}} f_2 z_2 dz_2 \ldots \int_a^b e^{xz_\mu\sqrt{-1}} f_\mu z_\mu dz_\mu .$$

[После преобразований]

$$X = Y e^{y\sqrt{-1}},$$

$\rho_1\rho_2\ldots\rho_\mu = Y$, $r_1 + r_2 + \ldots + r_\mu = y$.

Подставляя X в формулу (97.1), мы получаем […]

$$P = \frac{2}{\pi}\int_{-\infty}^{\infty} Y\cos(y - cx)\sin\varepsilon x\frac{dx}{x}. \qquad (101.1)$$

[…] При очень больших μ за исключением очень малых значений x произведение Y, равное единице при $x = 0$, становится, вообще говоря, очень малой дробью и окажется равным нулю, если μ может стать бесконечным. Не рассматривая, как и в § 95, особого случая[4], при котором Y стремится к величине, отличной от нуля, мы можем в интеграле формулы (101.1) придавать x лишь очень малые значения, вне которых Y оказывается неощутимым, так что

$Y = \exp(-\theta^2)$.

Поэтому, при замене переменной x на θ, пределами интеграла по θ следует считать нуль и бесконечность. Чтобы выразить x и dx через θ и $d\theta$, я разлагаю предыдущие выражения[5] $\rho_n\cos r_n$ и $\rho_n\sin r_n$ по степеням x и заменяю z_n на z под знаком интеграла. Полагая

$$\int_a^b zf_n zdz = k_n,\ \int_a^b z^2 f_n zdz = k'_n,\ \int_a^b z^3 f_n zdz = k''_n, \ldots,$$

мы получим сходящиеся ряды

$$\rho_n\cos r_n = 1 - \frac{x^2}{2!}k'_n + \frac{x^4}{4!}k'''_n - \ldots,\ \ \rho_n\sin r_n = xk_n - \frac{x^3}{3!}k''_n + \ldots$$

Далее, я ввожу

$(k'_n - k^2_n)/2 = h_n$, $(k''_n - 3k_n k'_n + 2k_n^3)/6 = g_n$, … […]

$\sum k_n = \mu k$, $\sum h_n = \mu h$, $\sum g_n = \mu g$, $\sum(l_n - h_n^2/2) = \mu l$, …

Здесь и ниже суммы распространяются от $n = 1$ до μ.

Имеет место формула

$$\ln Y = -\theta^2 = -x^2\mu h + x^4\mu l - \ldots,$$

откуда следует, что

$$x = \frac{\theta}{\sqrt{\mu h}} + \frac{l\theta^3}{2\mu h^2\sqrt{\mu h}} + \ldots,\ \frac{dx}{x} = \frac{d\theta}{\theta} + \frac{l\theta d\theta}{\mu h^2} + \ldots$$

[После преобразований оказывается, что]

$$P = \frac{2}{\pi}\int_0^{\infty} \exp(-\theta^2)\cos[(\mu k - c)x]\sin \varepsilon x \frac{d\theta}{\theta} +$$
$$\frac{2g}{\pi h\sqrt{\mu h}}\int_0^{\infty} \exp(-\theta^2)\sin[(\mu k - c)x]\sin \varepsilon x\ \theta^2 d\theta, \qquad (101.2)$$

и

$$P = 1 - \frac{2}{\sqrt{\pi}}\int_u^{\infty} \exp(-t^2)dt \qquad (101.3)$$

выражает вероятность, того, что при очень большом числе μ испытаний сумма *s* значений A заключена в границах

$$\mu k \mp 2u\sqrt{\mu h},\ k = c/\mu. \qquad (101.4)$$

При делении этой суммы на μ можно определить соответствующую вероятность для среднего значения *s*/μ.

102. Если даже *u* незначительно, вероятность (101.3) будет тем не менее весьма мало отличаться от единицы. Мы поэтому заключаем, что отношение *s*/μ вероятно очень немного отличается от *k*. Эта величина является суммой возможных значений A, умноженных на их шансы в каждом соответствующем испытании, и разделённой на число μ этих испытаний, т. е. суммой указанных значений, умноженных на их соответствующие средние шансы, и наше заключение совпадает с предложением § 53, которое таким образом доказано во всей общности.

Итак, при очень большом числе μ испытаний неизменно существует вероятность, весьма близкая достоверности, что среднее значение A очень немного отличается от *k*. Разность

($s/\mu - k$) неограниченно убывает с возрастанием μ и станет в точности нулем, если это число окажется бесконечным.

Если построить плоскую кривую в текущих координатах z и $f_n z$, она представит закон вероятностей значений A при n-м испытании, так что элемент $f_n z dz$ площади этой кривой будет бесконечно малой вероятностью значения A, выраженного абсциссой z. Кривая в текущих координатах z и $(1/\mu)\sum f_n z$ выразит закон вероятностей среднего значения A в серии μ испытаний. Как и полная площадь этой кривой от a до b, равная единице, интеграл

$$\int_a^b f_n z dz = 1. \qquad (102.1)$$

Если обозначить через ς абсциссу её центра тяжести, то

$$k = \frac{1}{\mu}\sum\int_a^b z f_n z dz = \varsigma.$$

Эта абсцисса и есть величина k, к которой в любых случаях стремится среднее значение A. Она равна нулю, если по природе A в каждом испытании её значения, равные по величине и противоположные по знаку, равновероятны, т. е. если $f_n(-z) = f_n z$ для всех значений n и z.

Константа h должна быть положительной. [Следует доказательство равенства]

$$4h_n = \int_a^b\int_a^b (z - z')^2 f_n z f_n z' dz dz'.$$

Величина $4h_n$ очевидно положительна и не может равняться нулю, потому что все элементы двойного интеграла положительны. То же будет иметь место для $\sum h_n$ и для h.

Простейшим будет случай, при котором в течение всей серии испытаний все возможные значения A равновероятны. При любом n будет тогда выполняться равенство $f_n z = 1/(b - a)$ и поэтому[6]

$$k_n = k = (a + b)/2,\ h_n = h = (a^2 + ab + b^2)/6 - (a + b)^2/8.$$

Границами s/μ, которым соответствует вероятность P, окажутся поэтому

$$\frac{1}{2}(a+b) \mp \frac{u(b-a)}{\sqrt{6\mu}}, \text{ или } \mp\frac{2ub}{\sqrt{6\mu}} \text{ если } a=-b. \quad (102.2a, b)$$

Пусть, к примеру (§ 82) $u = 0.4765$, тогда среднее s/μ с равной вероятностью окажется внутри или вне границ $0.389b/\sqrt{\mu}$. При $\mu = 600$ можно держать пари на равных, что s/μ не отклонится от нуля больше, чем на $0.4765b/3{\cdot}10 \approx 0.016b$.

Это тот случай, при котором в каждом испытании точка M попадает с равной вероятностью в любое положение на отрезке длиной $2b$. При очень большом числе μ испытаний среднее расстояние M от середины отрезка с вероятностью P не превзойдёт доли $2u/\sqrt{6\mu}$ от b. Если же точка M при каждом испытании попадает на круг радиуса b, причём равные расстояния M от его центра равновероятны, ясно, что вероятность f_nzdz расстояния z пропорциональна этому z. Если считать эту функцию постоянной при всех испытаниях, и заметить, что все возможные расстояния заключены между нулём и b, то, чтобы удовлетворить условию, подобному (102.1), надо будет принять $f_nz = 2z/b^2$. Поэтому[7]

$$k_n = k = 2b/3,\ 2h_n = 2h = b^2/2 - 4b^2/9$$

и P окажется вероятностью, что при μ испытаниях среднее расстояние M от центра будет заключаться в границах

$$\frac{2b}{3} \mp \frac{ub}{3\sqrt{\mu}}.$$

103. Мы (§ 97) предположили, что вещь A может принимать все, хоть и не равновероятные, значения между границами a и b, но выведенные нами формулы не менее применимы к случаю, при котором число возможных значений A ограничено. Для доказательства достаточно считать функции f_1z, f_2z, … которые выражают законы вероятностей значений A в μ последовательных испытаниях, разрывными[8].

Пусть $c_1, c_2, \ldots, c_\nu$ – ν значений z, заключённых между a и b, а функция f_nz не равна нулю при всех значениях z, бесконечно мало отличных от одной из величин $c_1, c_2, \ldots, c_v$. Обозначим через δ бесконечно малую и предположим также, что

$$\int_{c_1-\delta}^{c_1+\delta} f_n z dz = \gamma_1,\ \int_{c_2-\delta}^{c_2+\delta} f_n z dz = \gamma_2, \ldots,\ \int_{c_v-\delta}^{c_v+\delta} f_n z dz = \gamma_v.$$

Таким образом А будет принимать только v заданных значений c_1, $c_2, \ldots, c_v$, вероятности которых в n-м испытании равны $\gamma_1, \gamma_2, \ldots, \gamma_v$ и могут изменяться от одного испытания к другому, т. е. с числом n. Но поскольку одно из этих значений наверняка имеет место в n-м испытании, для всех значений n от 1 до μ должно выполняться равенство

$$\gamma_1 + \gamma_2 + \ldots + \gamma_v = 1.$$

Эта сумма кроме того равна значению интеграла от $f_n z$, а полученное уравнение заменяет условие (102.1). [После преобразований оказывается, что]

$$\int_a^b z f_n z dz = \gamma_1 c_1 + \gamma_2 c_2 + \ldots + \gamma_v c_v,$$
$$\int_a^b z^2 f_n z dz = \gamma_1 c_1^2 + \gamma_2 c_2^2 + \ldots + \gamma_v c_v^2,$$

и поэтому величины k и h из § 101 становятся равными

$$k = \frac{1}{\mu} \sum (\gamma_1 c_1 + \gamma_2 c_2 + \ldots + \gamma_v c_v),$$
$$h = \frac{1}{2\mu} \sum [(\gamma_1 c_1^2 + \gamma_2 c_2^2 + \ldots + \gamma_v c_v^2) - (\gamma_1 c_1 + \gamma_2 c_2 + \ldots + \gamma_v c_v)^2].$$

Суммирование здесь происходит по μ испытаниям и формула (101.3) выразит вероятность того, что сумма s значений А в этой серии испытаний будет заключена в границах (101.4), в которой за k и h должны быть взяты именно их найденные значения и которые легко вычислить, если для каждого испытания даны v возможных значений А и их соответствующие вероятности.

Если же эти вероятности постоянны и, более того, равны друг другу, их общее значение будет равно $1/v$ и тогда просто

$k = (c_1 + c_2 + \ldots + c_\nu)/\nu$,
$h = [\nu(c_1^2 + c_2^2 + \ldots + c_\nu^2) - (c_1 + c_2 + \ldots + c_\nu)^2]/2\nu^2$.

Пусть, к примеру, возможными значениями А будут шесть чисел, указанных на гранях обычной игральной кости, которую бросают последовательно очень большое число раз μ. Пренебрегая возможными небольшими различиями в шансах этих граней, мы имеем

$\nu = 6$, $c_1 = 1$, $c_2 = 2$, $c_3 = 3$, $c_4 = 4$, $c_5 = 5$, $c_6 = 6$,
$k = 7/2$, $h = 35/24$,

и формула (101.3) выразит вероятность того, что сумма s очков, вышедших при μ последовательных испытаниях, будет заключаться в границах

$$\frac{1}{2}\left(7\mu \mp u\sqrt{\frac{70\mu}{3}}\right).$$

Принимая $u = 0.4765$ и $\mu = 100$, сумма s будет с равной вероятностью заключаться в границах 350 ± 11.5 или вне их.

104. Рассмотрим теперь, как в § 52, событие Е некоторой природы, наступление которого может происходить только ввиду ν различных взаимоисключающих причин $C_1, C_2, \ldots, C_\nu$. Пусть причина C_i придаёт шанс c_i появлению события Е и γ_i – вероятность её действия. Поэтому шанс Е может изменяться от одного испытания к другому, принимая ν различных значений $c_1, c_2, \ldots, c_\nu$ при неизменных вероятностях $\gamma_1, \gamma_2, \ldots, \gamma_\nu$ и причинах $C_1, C_2, \ldots, C_\nu$. Если принять подобный шанс для Е, то будет существовать вероятность P (101.3) того, что при очень большом числе μ испытаний его среднее значение окажется в границах, определяемых формулой (101.4). Поскольку $c_1, c_2, \ldots, c_\nu$ и $\gamma_1, \gamma_2, \ldots, \gamma_\nu$ неизменны, величины k и h оказываются равными

$$k = \gamma_1 c_1 + \gamma_2 c_2 + \ldots + \gamma_\nu c_\nu, \qquad (104.1)$$
$$h = [(\gamma_1 c_1^2 + \gamma_2 c_2^2 + \ldots + \gamma_\nu c_\nu^2) - (\gamma_1 c_1 + \gamma_2 c_2 + \ldots + \gamma_\nu c_\nu)^2]/2$$

и, как усматривается, не зависимыми от μ, каково бы ни было число величин, включённых в них, и их неравенства друг другу.

Величине u можно придать небольшое значение, при котором окажется весьма близкая достоверности вероятность P того, что отличие среднего шанса события E в серии испытаний от суммы произведений в формуле (104.1) вероятно будет очень мало и станет неограниченно убывать по мере дальнейшего возрастания μ. В этом состояло второе из двух общих предложений § 52, которое нам оставалось доказать.

В двух сериях очень большого числа μ и μ' испытаний, в которых событие E наступало m и m' раз, отношения m/μ и m'/μ' вероятно очень немного отличаются от соответствующих средних шансов события E (§ 96). Поэтому весьма вероятно, что они очень мало отличаются от предыдущего значения k (104.1) и, стало быть, друг от друга, поскольку значения k в обеих сериях совпадают, если только причины C_1, C_2, … не изменились в промежутке времени между этими сериями. Но какова вероятность малой заданной разности между m/μ и m'/μ'? Этим важным вопросом мы займёмся ниже.

105. В большинстве вопросов, к которым применима формула (101.3), закон вероятностей значений A неизвестен, и величины k и h, включённые в границы среднего значения A, не определяются заранее. Но проведенная длинная серия испытаний может послужить для исключения неизвестных из границ среднего значения A из другой длинной серии испытаний, подверженных действию тех же причин, которые придают тот же шанс каждому значению A и сами имеют ту же вероятность. Полное решение этой задачи является предметом проведенных ниже вычислений.

Я принимаю, что в формуле (101.2) $c = \varepsilon$ […] и определяю при этом условии вероятность сумме s значений A в μ испытаниях заключаться в границах от нуля до 2ε. Далее, производная от P по ε […] выразит бесконечно малую вероятность того, что в точности $s = 2\varepsilon$. Пусть также

$$2\varepsilon = \mu k + 2\nu\sqrt{\mu h},\ d\varepsilon = \sqrt{\mu h}d\nu.$$

Обозначим через wdv соответствующее значение $(dP/d\varepsilon)d\varepsilon$, в котором величинами порядка малости $1/\mu$ можно пренебрегать, что позволит свести x к первому члену его выражения в § 101, т. е. к $\theta/\sqrt{\mu h}$. Тогда [после преобразований]

$$\frac{s}{\mu} = k + \frac{2v\sqrt{h}}{\sqrt{\mu}},$$

где[9] v положительно или отрицательно, но очень мало по сравнению с $\sqrt{\mu}$. Я обозначу через $C_1, C_2, \ldots, C_v$ все взаимоисключающие причины, известные или неизвестные, которые могут придать вещи A одно из её возможных значений, и через $\gamma_1, \gamma_2, \ldots, \gamma_v$ их соответствующие вероятности, в сумме составляющие единицу. Если число причин бесконечно, то каждая из них бесконечно мала. Все значения A заключены между a и b; если их число бесконечно, шанс каждого, вызванный каждой из указанных причин, будет бесконечно мал. Я обозначу через $Z_i dz$ шанс, который доставляет причина C_i, если она достоверна, значению z вещи A. Интеграл

$$\int_a^b z f_n z dz, \qquad (105.1)$$

относящийся к n-му испытанию, может иметь v значений, соответствующих интегралам от $zZ_1, zZ_2, \ldots, zZ_v$ в тех же пределах, вероятностями которых являются вероятности их причин. Таким образом, для некоторого испытания γ_i выражает шанс значения интеграла от функции zZ_i. И поэтому бесконечно малая вероятность среднего значения

$$\frac{1}{\mu}\sum\int_a^b z f_n z dz$$

определяется по предыдущему правилу о среднем значении s/μ некоторой вещи из очень большого числа μ испытаний: s есть сумма μ неизвестных значений интеграла (105.1) в этой серии испытаний, а величины k и h должны определяться по его v возможным значениям.

Примем эти v значений с соответствующими Z_i вместо $c_1, c_2, \ldots, c_v$ (§ 103), и положим для сокращения письма

$$\gamma = \sum \gamma_i \int_a^b z Z_i dz,$$

$$\beta = \frac{1}{2}\sum \gamma_i \left(\int_a^b zZ_i dz \right)^2 - \frac{1}{2}\left(\sum \gamma_i \int_a^b zZ_i dz \right)^2, \qquad (105.2)$$

где[10] суммы распространяются от $i = 1$ до v. В соответствии с формулами того параграфа, эти γ и β, независимые от μ, и должны быть приняты за k и h. Обозначим через v_1 положительную или отрицательную величину, очень малую по сравнению с $\sqrt{\mu}$, и через V_1 – многочлен, содержащий лишь нечётные степени v_1, и введём бесконечно малую величину $w_1 dv_1$

$$w_1 dv_1 = \frac{1}{\sqrt{\pi}}[1 - \frac{V_1}{\sqrt{\mu}}]\exp(-v_1^2)dv_1,$$

которая оказывается вероятностью уравнения

$$\frac{1}{\mu}\sum \int_a^b zf_n z dz = \gamma + \frac{2v_1\sqrt{\beta}}{\sqrt{\mu}}. \qquad (105.3)$$

Рассмотрим также выражение

$$\frac{1}{2}\int_a^b z^2 f_n z dz - \frac{1}{2}\left(\int_a^b z f_n z dz \right)^2$$

как вещь, способную принимать v значений в соответствии с причинами $C_1, C_2, \ldots, C_v$ с вероятностями в каждом испытании, совпадающими с вероятностями этих причин. Кроме того, обозначим через $\hat{v}$ такую положительную или отрицательную величину, что $\hat{v}/\sqrt{\mu}$ окажется очень малой дробью, и через $\hat{V}$ – многочлен, содержащий лишь нечётные степени $\hat{v}$, введём величину

$$\hat{w}d\hat{v} = \frac{1}{\sqrt{\pi}}[1 - \frac{\hat{V}}{\sqrt{\mu}}]\exp(-\hat{v}^2)d\hat{v}$$

и положим для сокращения письма

$$\alpha = \frac{1}{2}\sum \gamma_i \int_a^b z^2 Z_i dz - \frac{1}{2}\sum \gamma_i \left(\int_a^b zZ_i dz \right)^2.$$

Выражение $\widehat{w}d\widehat{v}$ окажется вероятностью того, что среднее из μ значений интересующей нас величины, а именно

$$\frac{1}{2\mu}\sum\int_a^b z^2 f_n z dz - \frac{1}{2\mu}\sum\left(\int_a^b z f_n z dz\right)^2,$$

будет отличаться от α лишь на определённую величину порядка малости $1/\sqrt{\mu}$, устанавливать которую нет смысла. Впрочем, это среднее просто совпадает с величиной *h* из § 101. И, если пренебрегать величинами порядка $1/\mu$, во втором члене предыдущего значения s/μ, который уже имеет порядок $1/\sqrt{\mu}$, будет достаточно принять α вместо *h.* Таким образом,

$$\frac{s}{\mu} = k + \frac{2v\sqrt{\alpha}}{\sqrt{\mu}},$$

причём вероятность этого равенства, если только принимаемое значение *h* достоверно, снова равна wdv. Но это значение имеет лишь вероятность $\widehat{w}d\widehat{v}$, зависящую от переменной $\widehat{v}$, которая не входит в значение s/μ. Поэтому полная вероятность этого последнего значения равна произведению wdv и суммы значений $\widehat{w}d\widehat{v}$, соответствующих всем возможным значениям $\widehat{v}$. Хотя эти значения должны быть очень малы по сравнению с $\sqrt{\mu}$, ввиду экспоненциальной функции $\exp(-\widehat{v}^2)$ в сомножителе $\widehat{w}d\widehat{v}$ можно без его ощутимого изменения распространить интеграл от него от $\widehat{v} = -\infty$ до ∞. Часть, зависящая от $\widehat{V}$, исчезнет, поскольку она состоит из элементов, попарно равных по величине и противоположных по знаку, так что просто

$$\int_{-\infty}^{\infty} \overline{w}d\widehat{v} = 1.$$

Следовательно, вероятность предыдущего уравнения будет неизменно равна wdv, как будто бы принятое нами приближённое значение *h* являлось достоверным. Можно также заметить, что среднее, составляющее левую часть уравнения (105.3), совпадает с величиной *k* из § 101, а w_1dv_1 является вероятностью того, что эта величина равна

$$k = \gamma + \frac{2v_1\sqrt{\beta}}{\sqrt{\mu}}.$$

Подставляя это значение в s/μ, мы получим

$$\frac{s}{\mu} = \gamma + \frac{2v_1\sqrt{\beta}}{\sqrt{\mu}} + \frac{2v\sqrt{\alpha}}{\sqrt{\mu}},$$

т. е. вероятность того, что это последнее уравнение для каждой пары значений v и v_1 оказывается произведением $wdvw_1dv_1$, которое, пренебрегая членом с μ в знаменателе, я обозначу через σ:

$$\sigma = \frac{1}{\pi}[1 - \frac{1}{\sqrt{\mu}}(V + V_1)]\exp(-v^2 - v_1{}^2)dvdv_1.$$

Обозначим через θ положительную или отрицательную величину, очень небольшую, как и v и v_1, по сравнению с $\sqrt{\mu}$. Тогда можно будет полагать, что

$$v_1\sqrt{\beta} + v\sqrt{\alpha} = \theta\sqrt{\alpha + \beta}.$$

Если желательно, в предыдущей дифференциальной формуле можно заменить v_1 и dv_1 этой новой переменной:

$$v_1 = \frac{\sqrt{\alpha + \beta}}{\sqrt{\beta}} - \frac{v\sqrt{\alpha}}{\sqrt{\beta}},\ dv_1 = \frac{\sqrt{\alpha + \beta}}{\sqrt{\beta}}d\theta.$$

[После преобразований оказалось, что уравнение]

$$\frac{s}{\mu} = \gamma + \frac{2\theta\sqrt{\alpha + \beta}}{\sqrt{\mu}} \qquad (105.4)$$

содержит лишь переменную θ и поэтому его полная вероятность будет равна сумме значений σ при всех положительных и отрицательных значениях, которые могут быть приданы другой переменной, v. Более того, ввиду экспоненциальной функции, содержащейся в выражении для σ, окажется допустимым распространить этот интеграл [?] от $v = -\infty$ до ∞ без ощутимого изменения его значения.

Итак, […]

$$\eta d\theta = \frac{1}{\sqrt{\pi}}\exp(-\theta^2)d\theta - \frac{1}{\sqrt{\pi\mu}}\chi\exp(-\theta^2)d\theta$$

есть вероятность уравнения (105.4); χ – многочлен, содержащий лишь нечётные степени θ. Требуется исключить из этого уравнения неизвестное (α + β). Это возможно, потому что оно сводится к

$$\alpha + \beta = \frac{1}{2}\sum\gamma_i\int_a^b z^2 Z_i dz - \frac{1}{2}\left(\sum\gamma_i\int_a^b zZ_i dz\right)^2 \qquad (105.5)$$

и независимо от суммы $\sum\gamma_i\left(\int_a^b zZ_i dz\right)^2$, которая содержится и в α, и в β.

106. [Автор вводит обозначение]

$$\varphi = \frac{1}{\mu}\sum\int_a^b z^2 f_n z dz$$

[и доказывает, что] существует вероятность $\widehat{w}d\widehat{v}$ того, что

$$\frac{1}{2}\sum\gamma_i\int_a^b z^2 Z dz$$

отличается от φ/2 лишь на определённую величину порядка малости 1/√μ. Более того, если неизменно пренебрегать членами, содержащими μ в знаменателе, то оказывается, что, как в § 105, позволительно заменить первое слагаемое выражения α + β на φ/2, нисколько не меняя вероятности $\eta d\theta$ этого равенства. Другая часть значения α + β в точности равна $\gamma^2/2$, так что

$$\alpha + \beta = \varphi/2 - \gamma^2/2,$$

а уравнение (105.4) принимает вид

$$\frac{s}{\mu}=\gamma+\frac{\theta\sqrt{2\varphi-2\gamma^2}}{\sqrt{\mu}}.$$

Пусть теперь Z является данной функцией z. Анализ в §§ 97 и 101, а затем выражение wdv из § 105 без затруднений распространяются на сумму значений Z в μ рассматриваемых нами испытаний, для чего достаточно заменить А другой вещью A_1 со значениями этой функции Z. Бесконечно малая вероятность какого-либо значения A_1 окажется той же, что у соответствующего значения z, и поэтому в n-м испытании она равна $f_n z dz$. Если обозначить через k_1, h_1, g_1, … величины, относящиеся теперь к A_1, которые в § 101 относились к А и обозначались через k, h, g, …, то

$$\frac{s_1}{\mu}=k_1+\frac{2v\sqrt{h_1}}{\sqrt{\mu}}$$

и если $Z = z^2$,

$$k_1=\frac{1}{\mu}\sum\int_a^b z^2 f_n z dz=\varphi.$$

В предыдущем выражении s/μ с принятой нами степенью приближения $\varphi = s_1/\mu$. Как и в § 105, мы удостоверимся в том, что вероятность этого выражения не изменится и $\eta d\theta$ остаётся бесконечно малой вероятностью уравнения

$$\frac{s}{\mu}=\gamma+\frac{\theta\sqrt{2s_1/\mu-2\gamma^2}}{\sqrt{\mu}}\ \text{ или }\ \frac{s}{\mu}=\gamma+\frac{\theta\sqrt{2s_1/\mu-2s^2/\mu^2}}{\sqrt{\mu}},$$

что можно вывести из предыдущего, если снова пренебрегать величинами порядка малости $1/\mu$.

Я обозначу через λ_n значение А при n-м испытании и для сокращения письма введу

$$\sum\lambda_n/\mu = \lambda,\ \sum(\lambda_n-\lambda)^2/\mu = l^2/2.$$

Тождественно выполняются равенства

$$\frac{s_1}{\mu}=\frac{\sum\lambda_n^2}{\mu},\ \frac{s}{\mu}=\frac{\sum\lambda_n}{\mu},\ \frac{s_1}{\mu}-\frac{s}{\mu}=\frac{\sum(\lambda_n-\lambda)^2}{\mu}$$

и предыдущее уравнение принимает вид

$$\frac{s}{\mu}=\gamma+\frac{\theta l}{\sqrt{\mu}}.$$

Поэтому, если обозначить через u заданную положительную величину, интеграл от вероятности $\eta d\theta$ этого уравнения в пределах $\theta = -u$ и u выразит вероятность того, что s/μ окажется в границах $\gamma \pm ul/\sqrt{\mu}$. Обозначив эту вероятность через Γ и учитывая выражение $\eta d\theta$, мы получим

$$\Gamma=\frac{1}{\sqrt{\pi}}\int_{-u}^{u}\exp(-\theta^2)d\theta-\frac{1}{\sqrt{\pi\mu}}\int_{-u}^{u}\chi\exp(-\theta^2)d\theta.$$

Но так как χ – это многочлен, содержащий только нечётные степени θ, второй интеграл равен нулю, и Γ, как оказывается, совпадает с вероятностью P в формуле (101.3).

Итак, эта формула выражает вероятность того, что границы $\pm\, ul/\sqrt{\mu}$, которые после испытаний уже не зависят ни от каких неизвестных, содержит разность между средним s/μ значением А и особой величиной γ. Это среднее неограниченно приближается к γ и достигнет её, если μ станет бесконечным, но причины C_1, C_2, …, C_v возможных значений А оставались бы при этом неизменными.

107. Пусть произведено две серии большого числа испытаний, μ и μ'. Пусть суммы значений А в этих сериях обозначены через s и s', а λ_n и λ'_n пусть будут значениями А в n-х испытаниях. Примем

$$\sum\lambda_n/\mu = \lambda,\ \sum(\lambda_n-\lambda)^2/\mu = l^2/2,\ \sum\lambda'_n/\mu' = \lambda',\ \sum(\lambda'_n-\lambda')^2/\mu' = l'^2/2.$$

Суммы распространяются на все испытания каждой серии, т. е. в первой серии от $n = 1$ до μ, а во второй серии – от $n = 1$ до μ'. Если между этими сериями причины C_1, C_2, …, C_v не изменились, величина γ (105.2) также останется без изменения. Обозначим через θ и θ' положительные или отрицательные переменные, очень малые по сравнению с $\sqrt{\mu}$ и $\sqrt{\mu'}$. Уравнения, относящиеся к средним значениям А в этих двух сериях, будут иметь вид[11]

$$\frac{s}{\mu} = \gamma + \frac{\theta l}{\sqrt{\mu}}, \quad \frac{s'}{\mu'} = \gamma + \frac{\theta' l'}{\sqrt{\mu'}}, \qquad (107.1)$$

а их вероятности $\eta d\theta$ и $\eta' d\theta'$ окажутся равными

$$\eta d\theta = \frac{1}{\sqrt{\pi}}[1 - \frac{\chi}{\sqrt{\mu}}]\exp(-\theta^2)d\theta,$$
$$\eta' d\theta' = \frac{1}{\sqrt{\pi}}[1 - \frac{\chi'}{\sqrt{\mu'}}]\exp(-\theta'^2)d\theta'.$$

Здесь χ и χ' – многочлены, включающие только нечётные степени θ и θ'. Далее, если серии различаются по своим испытаниям, можно считать s/μ и s'/μ' независимыми событиями. По правилу § 5 вероятность их совместного наступления равна произведению $\eta d\theta$ и $\eta' d\theta'$. Это же имеет место для вероятности каких-либо комбинаций уравнений (107.1) и в частности для их разности. […] Итак, пренебрегая членами с $\sqrt{\mu\mu'}$ в знаменателе, мы получаем вероятность указанной разности для каждой пары значений θ и θ'

$$\psi \equiv \eta\eta' d\theta d\theta' = \frac{1}{\pi}[1 - \frac{\chi}{\sqrt{\mu}} - \frac{\chi'}{\sqrt{\mu'}}]\exp(-\theta^2 - \theta'^2)d\theta d\theta'. \qquad (107.2)$$

Поступая как в § 105, я принимаю

$$\frac{\theta' l'}{\sqrt{\mu'}} - \frac{\theta l}{\sqrt{\mu}} = \frac{t\sqrt{l'^2\mu + l^2\mu'}}{\sqrt{\mu\mu'}},$$

так что

$$\frac{s'}{\mu'} - \frac{s}{\mu} = \frac{t\sqrt{l'^2\mu + l^2\mu'}}{\sqrt{\mu\mu'}}.$$

Далее, я заменяю θ в уравнении (107.2) новой переменной t, и с этой целью принимаю, что

$$\theta' = \frac{t\sqrt{l'^2\mu + l^2\mu'}}{l'\sqrt{\mu}} + \frac{\theta l\sqrt{\mu'}}{l'\sqrt{\mu}},\ d\theta' = \frac{\sqrt{l'^2\mu + l^2\mu'}}{l'\sqrt{\mu}}dt$$

и поэтому

$$\psi = \frac{dtd\theta\sqrt{l'^2\mu + l^2\mu'}}{\pi l'\sqrt{\mu}}(1-\Pi)\exp(-\theta'^2 - t^3)\,,$$

где Π – многочлен, каждый член которого содержит нечётную степень t или θ. Разность $s'/\mu' - s/\mu$ содержит лишь переменную t, а её вероятность будет равна интегралу от ψ, распространённому на все значения, которые может принимать другая переменная, θ. Ввиду экспоненциальной функции, включённой в ψ, этот интеграл может без ощутимого изменения его значения рассматриваться в бесконечных пределах. Принимая поэтому $\theta' = t'$, $d\theta' = dt'$ и обозначая через Π' соответственно изменённое Π, мы получаем

$$\psi = \frac{1}{\pi}(1-\Pi')\exp(-t'^2 - t^2)dt'dt\,.$$

[После преобразований оказывается, что]

$$\mp\frac{u\sqrt{l'^2\mu + l^2\mu'}}{\sqrt{\mu\mu'}}$$

являются границами, внутри которых с вероятностью

$$\Delta = \frac{2}{\sqrt{\pi}}\int_0^u \exp(-t^2)dt$$

находится разность $s'/\mu' - s/\mu$. Эта вероятность совпадает со значением P в формуле (101.3). Таким образом, P является вероятностью того, что разность средних значений A в двух длинных сериях испытаний будет заключена в границах, не содержащих ничего неизвестного.

Приняв для u значение, достаточное для того, чтобы P очень мало отличалось от единицы, и обнаружив, что указанная разность оказывается вне установленных границ, мы обоснованно можем заключить, что между сериями испытаний причины C_1, C_2, …, C_ν возможных значений А не оставались неизменными, т. е. что произошло некоторое изменение вероятностей γ_1, γ_2, …, γ_ν этих причин или шансов, которые они придают различным значениям А.

В соответствии со сказанным в § 106, каждая из величин l и l' весьма вероятно должна очень мало отличаться от неизвестной и постоянной в обеих сериях испытаний величины $2\sqrt{\alpha+\beta}$. И поэтому весьма вероятно, что они очень мало отличаются друг от друга. Не изменяя ощутимо ни указанных выше границ, ни их вероятности, можно поэтому считать, что $l' = l$. В будущей серии испытаний будет иметь место определяемая по формуле (101.3) вероятность P того, что границами среднего s'/μ' значения А служат

$$\frac{s}{\mu} \mp \frac{ul\sqrt{\mu+\mu'}}{\sqrt{\mu\mu'}},$$

которые при каждом заданном значении u зависят лишь от результатов уже произведенной первой серии испытаний.

При одном и том же значении u, т. е. с равной степенью вероятности, интервал между этими границами больше, чем для разности $\gamma - s/\mu$, в отношении $\sqrt{\mu+\mu'}$ к $\sqrt{\mu'}$. Интервалы для обеих серий почти совпадают, если μ' очень большое число по сравнению с очень большим числом μ.

108. Если две серии μ и μ' испытаний произведены для измерения одной и той же вещи различными инструментами, у которых равные по величине и противоположные по знаку погрешности равновероятны, средние значения s/μ и s'/μ', выведенные из этих серий, будут неограниченно стремиться к одной и той же величине, которая окажется истинным значением А (§ 60)[12]. В этом случае неизвестное γ будет, следовательно, одним и тем же в обеих сериях наблюдений, а средние s/μ и s'/μ' весьма вероятно окажутся лишь очень немного различными. Тем не менее, значения неизвестного $\alpha + \beta$ могут очень намного отличаться друг от друга, а потому весьма неравными будут l и l'. Эти величины известны, и можно выяснить, как лучше всего комбинировать средние s/μ и s'/μ' для установления границ γ или истинного значения A.

Для этой цели я обозначу через g и g' неопределённые величины, в сумме составляющие единицу, и сложу уравнения (107.1), предварительно умножив их соответственно на g и g':

$$\gamma = \frac{gs}{\mu} + \frac{g's'}{\mu'} - \frac{gl\theta}{\sqrt{\mu}} - \frac{g'l'\theta'}{\sqrt{\mu'}}.$$

В соответствии со сказанным выше, при любой паре значений θ и θ' вероятность полученного уравнения равна ψ и из вычислений, подобных только что выполненным, следует, что величина P, данная формулой (101.3), выражает вероятность того, что неизвестное γ заключено в границы

$$\frac{gs}{\mu} + \frac{g's'}{\mu'} \mp \frac{u\sqrt{g'^2l'^2\mu + g^2l^2\mu'}}{\sqrt{\mu\mu'}}.$$

Если желательно, чтобы при одной и той же вероятности P, т. е. для каждого заданного значения u, интервал между этими границами был как можно короче, нужно определять g и g', приравнивая нулю производную коэффициента при u по этим величинам. Ввиду равенств $g + g' = 1$ и $dg' = -dg$ оказывается, что

$$g = \frac{l'^2\mu}{l'^2\mu + l^2\mu'},\ g' = \frac{l^2\mu'}{l'^2\mu + l^2\mu'},$$

а самыми тесными границами γ будут

$$\frac{sl'^2 + s'l^2}{l'^2\mu + l^2\mu'} \mp \frac{ull'}{\sqrt{l'^2\mu + l^2\mu'}},$$

вероятность которых неизменно указывает формула (101.3).

Этот результат можно легко обобщить на любое число серий из многочисленных наблюдений, произведенных различными инструментами для измерения одной и той же вещи A. Пусть величины μ, s, l в первой серии будут обозначены во второй и в третьей сериях соответственно через μ', s', l' и μ'', s'', l'' и т. д. Примем также, что

$$\frac{\mu}{l^2}+\frac{\mu'}{l'^2}+\frac{\mu''}{l''^2}+\ldots=D^2,\ \frac{\mu}{D^2l^2}=q,\ \frac{\mu'}{D^2l'^2}=q',\ \frac{\mu''}{D^2l''^2}=q'',\ \ldots$$

Тогда формула (101.3) выразит вероятность того, что неизвестное значение A будет заключено в границах

$$\frac{sq}{\mu}+\frac{s'q'}{\mu'}+\frac{s''q''}{\mu''}+\ldots\mp\frac{u}{D},$$

полученных самой благоприятной комбинацией наблюдений. Принимая незначительное значение для u, величину P в формуле (101.3) можно сделать очень близкой к единице, и потому значение A весьма вероятно будет очень немного отличаться от суммы средних s/μ, s'/μ', s''/μ'', …, умноженных соответственно на величины q, q', q'', … Результат каждой серии наблюдений тем больше влияет на это приближённое значение A и на интервал между его границами $\mp u/D$, чем значительнее соответствующие ей отношения μ/l^2, или μ'/l'^2, или μ''/l''^2, …

Если все серии наблюдений были выполнены одним и тем же инструментом, их можно будет считать единой серией из $\mu + \mu' + \mu'' + \ldots$ наблюдений. Как сказано выше, величины l, l', l'', … будут весьма вероятно почти равны друг другу. Полагая суммы распространёнными на всю общую серию от $n = 1$ до $\mu + \mu' + \mu'' + \ldots$ и вводя

$$\frac{\sum\lambda_n}{\mu+\mu'+\mu''+\ldots}=\lambda,\ \frac{\sum(\lambda_n-\lambda)^2}{\mu+\mu'+\mu''+\ldots}=\frac{l_1^2}{2},$$

мы можем, как сказано выше, принять l_1 за общее значение l, l', l'', … Предыдущие границы неизвестного γ, вероятность которых указывает формула (101.3), становятся равными

$$\frac{s+s'+s''+\ldots}{\mu+\mu'+\mu''+\ldots}\mp\frac{ul_1}{\sqrt{\mu+\mu'+\mu''+\ldots}},$$

что совпадает с результатом § 106, относящимся к одной-единственной серии испытаний.

109. Вопрос, поставленный в конце § 104, разрешается методом, подобным тому, который был только что применен. Пусть в очень большом числе μ испытаний событие E какой-либо природы произошло m раз. Шанс E изменяется и равен p_n при n-м испытании. Пусть $\sum p_n/\mu = p$ и $\sum p_n^2 = q$. Обозначим через v положительную или отрицательную величину, очень малую по сравнению с $\sqrt{\mu}$, а U пусть будет вероятностью уравнения

$$\frac{m}{\mu} = p - \frac{v\sqrt{2p-2q}}{\sqrt{\mu}}.$$

Пренебрежём для упрощения вычислений вторым слагаемым формулы (95.1), учтём значение включённой туда величины k и заменим θ на v, тогда

$$U = \frac{\exp(-v^2)}{\sqrt{2\pi\mu(p-q)}}.$$

Как и в § 104, обозначим все возможные причины события E, число которых может быть конечно или бесконечно, через C_1, C_2, …, C_v, причём $\gamma_1, \gamma_2, \ldots, \gamma_v$ будут их вероятностями, а $c_1, c_2, \ldots, c_v$ – шансами, которые они придают наступлению E. Полагая, что p_n может принимать эти v значения, вероятности которых равны γ_1, $\gamma_2, \ldots, \gamma_v$, положим, что

$$r = \gamma_1 c_1 + \gamma_2 c_2 + \ldots + \gamma_v c_v,\ \rho = \gamma_1 c_1^2 + \gamma_2 c_2^2 + \ldots + \gamma_v c_v^2.$$

Обозначим через v_1 положительную или отрицательную переменную, очень малую по сравнению с $\sqrt{\mu}$. Тогда бесконечно малая вероятность точного равенства

$$p = r + \frac{v_1\sqrt{2\rho - 2r^2}}{\sqrt{\mu}} \qquad (109.1)$$

окажется равной $w_1 dv_1$ (§ 105) или просто $(1/\sqrt{\pi})\exp(-v_1^2)dv_1$, если пренебречь вторым членом его выражения. Обозначим кроме того через $\hat{v}$ очень малую по сравнению с $\sqrt{\mu}$ переменную. Тогда вероятность того, что $p - q$ будет отличаться от $r - \rho$ лишь на определённую величину, пропорциональную $\hat{v}$ порядка малости $1/\sqrt{\mu}$, и будет равна $\widehat{w}d\widehat{v}$ или просто $(1/\sqrt{\pi})\exp(-\widehat{v}^2)d\widehat{v}$, см. § 105.

Далее, заметно, что, пренебрегая величинами порядка $1/\mu$, можно, не изменяя вероятности U значения предыдущего отношения m/μ, заменить $p - q$ на $r - \rho$. Соответственно, это отношение примет вид

$$\frac{m}{\mu} = p - \frac{v\sqrt{2r-2\rho}}{\sqrt{\mu}}. \qquad (109.2)$$

Введём теперь величину

$$\delta = \frac{1}{\sqrt{2\mu(r-\rho)}}.$$

Поскольку m – целое число, v может быть только положительным или отрицательным кратным δ, которое весьма мало по сравнению с $1/\sqrt{\mu}$.

Из формул (109.1) и (109.2) следует, что

$$\frac{m}{\mu} = r + \frac{v_1\sqrt{2\rho-2r^2}}{\sqrt{\mu}} - \frac{v\sqrt{2r-2\rho}}{\sqrt{\mu}}.$$

Вероятность этого уравнения для любой пары значений v и v_1 равна произведению U и $(1/\sqrt{\pi})\exp(-v_1^2)dv_1$, и я обозначу её через ε. При замене $p - q$ на $r - \rho$ в выражении для U это произведение оказывается равным

$$\varepsilon = \frac{1}{\pi\sqrt{2\mu(r-\rho)}}\exp(-v^2 - v_1^2)dv_1.$$

Пусть

$$v_1 = \theta\sqrt{\frac{r-r^2}{\rho-r^2}} + v\sqrt{\frac{r-\rho}{\rho-r^2}},\ dv_1 = \sqrt{\frac{r-r^2}{\rho-r^2}}d\theta,$$

тогда, если пренебречь членами порядка малости $1/\mu$,

$$\frac{m}{\mu} = r + \frac{\theta\sqrt{2r - 2r^2}}{\sqrt{\mu}},\ r = \frac{m}{\mu} - \frac{\theta\sqrt{2m(\mu - m)}}{\mu\sqrt{\mu}}.$$

[Далее автор выписывает формулу для ε, введя в неё δ и θ.]

Но выражение для r не содержится в v и также не зависит от этой величины. Оно равно сумме значений ε, соответствующих всем тем, которые может принимать v, и она должна возрастать по приращениям, равным δ, кратным которого является v. Ввиду малости δ можно вывести приближённое значение этой суммы, подставив dv вместо δ в ε и заменяя сумму интегралом. Полученное значение будет точным до величин порядка δ или $1/\sqrt{\mu}$. Хотя переменная v должна быть очень малой по сравнению с $\sqrt{\mu}$, ввиду экспоненциальной функции, содержащейся в ε, можно, не изменяя ощутимо его величины, распространить интеграл от
$v = -\infty$ до ∞. И если принять

$$v\sqrt{\frac{r - r^2}{\rho - r^2}} + \theta\sqrt{\frac{r - \rho}{\rho - r^2}} = \theta_1,\ d\theta_1 = \sqrt{\frac{r - r^2}{\rho - r^2}}dv,$$

пределы интеграла по θ_1 будут также бесконечны. Обозначая теперь через $\chi d\theta$ бесконечно малую вероятность выражения для r, мы получим

$$\chi d\theta = \frac{d\theta}{\pi}\exp(-\theta^2)\int_{-\infty}^{\infty}\exp(-\theta_1^2)d\theta_1 = \frac{1}{\sqrt{\pi}}\exp(-\theta^2)d\theta.$$

Итак, u – положительная и заданная величина, и вероятность того, что неизвестное значение r окажется в границах

$$\frac{m}{\mu} \mp \frac{u\sqrt{2m(\mu - m)}}{\mu\sqrt{\mu}} \tag{109.3}$$

совпадает с величиной P из формулы (101.3), потому что эта вероятность равна

$$\int_{-u}^{u}\chi d\theta = \frac{2}{\sqrt{\pi}}\int_{0}^{u}\exp(-\theta^2)d\theta.$$

Таким образом, P есть вероятность того, что особая величина r, к которой неограниченно стремится отношение m/μ по мере дальнейшего возрастания большого числа μ, не отличается от этого отношения больше, чем на величину, содержащуюся в границах, указанных формулой (109.3) и не включающих ничего неизвестного.

Пусть во второй серии из очень большого числа μ' испытаний событие Е появилось m' раз. Обозначим через θ' положительную или отрицательную переменную, очень малую по сравнению с $\sqrt{\mu'}$. Бесконечно малая вероятность уравнения

$$r = \frac{m'}{\mu'} - \frac{\theta'\sqrt{2m'(\mu' - m')}}{\mu'\sqrt{\mu'}}$$

будет равна $(1/\sqrt{\pi})\exp(-\theta'^2)d\theta'$, а вероятность уравнения

$$\frac{m'}{\mu'} - \frac{m}{\mu} = \frac{\theta'\sqrt{2m'(\mu' - m')}}{\mu'\sqrt{\mu'}} - \frac{\theta\sqrt{2m(\mu - m)}}{\mu\sqrt{\mu}},$$

полученного вычитанием этого значения r из его предыдущего значения, равна произведению $(1/\sqrt{\pi})\exp(-\theta'^2)d\theta'$ и $(1/\sqrt{\pi})\exp(-\theta^2)d\theta$ для каждой пары значений θ и θ' [значений θ' и θ].

[Автор заменяет] переменную θ' на t не изменяя θ, затем θ на t' не изменяя t. Вероятность предыдущего уравнения становится равной $(1/\pi)\exp(-t^2 - t'^2)dtdt'$, само же уравнение принимает вид

$$\frac{m'}{\mu'} - \frac{m}{\mu} = \frac{t\sqrt{2\mu^3 m'(\mu' - m') + 2\mu'^3 m(\mu - m)}}{\mu\mu'\sqrt{\mu\mu'}}.$$

Оно теперь содержит лишь переменную t, а его полная вероятность оказывается равной интегралу по t' от приведенного выше дифференциального выражения. Без ощутимого изменения его значения этот интеграл можно распространить от $t' = -\infty$ до ∞, что приведёт указанное выражение к $(1/\sqrt{\pi})\exp(-t^2)dt$ и потому к величине P из формулы (101.3). Она выражает вероятность, что разность $m'/\mu' - m/\mu$ заключена в границах
`

$$\mp \frac{u\sqrt{2\mu^3 m'(\mu'-m')+2\mu'^3 m(\mu-m)}}{\mu\mu'\sqrt{\mu\mu'}}.$$

Здесь u положительно и задано, и формула содержит только известные числа. Выведенные границы совпадают с полученными в § 87 гораздо проще, хотя лишь для случая, при котором шанс события Е постоянен и одинаков в обеих сериях испытаний. В формуле (87.2) содержится член порядка $1/\sqrt{\mu}$ или $1/\sqrt{\mu'}$, которого нет в формуле (101.3), потому что мы пренебрегли членами рассмотренных вероятностей этого порядка малости.

110. В этом сочинении я не предполагаю исследовать многочисленные вопросы, к которым можно приложить предыдущие формулы; основные из них были указаны в § 60 и последующих[13]. В качестве примера я приведу известный вопрос об орбитах планет и комет. [После длительных преобразований оказалось, что выведенная Пуассоном формула]

$$P = \frac{1}{\mu!}[\beta^\mu - \mu(\beta-1)^\mu + C_\mu^2(\beta-2)^\mu - C_\mu^3(\beta-3)^\mu + \ldots$$
$$- \alpha^\mu + \mu(\alpha-1)^\mu - C_\mu^2(\alpha-2)^\mu + C_\mu^3(\alpha-3)^\mu - \ldots], \quad (110.1)$$

совпадёт с формулой Лапласа (Laplace 1812/1886, p. 261), полученной с той же целью совсем другим путём. [Величины α и β определяются из равенств, введенных автором выше:]

$$h = g,\ \gamma = \mu g - c,\ c - \varepsilon = 2g\alpha,\ c + \varepsilon = 2g\beta.$$

Если все значения А в границах от нуля до $2g$ равно возможны и невозможны вне их, то формула (110.1) выразит вероятность того, что в некотором числе μ испытаний сумма значений какой-либо вещи А будет заключена между $2g\alpha$ и $2g\beta$. Ряды, содержащие в этой формуле β и α, продолжаются до тех пор, пока величины, возвышаемые в степень μ, не перестанут быть положительными. Если n представляет наибольшее целое число, содержащееся в β, то соответствующий ряд останавливается на $(n + 1)$-м члене или раньше в зависимости от неравенств $\mu > n$ или $< n$. То же имеет место для второго ряда для наибольшего целого числа, содержащегося в α.

Какова бы ни была причина, определившая образование планет, предполагают, что вначале все возможные наклоны плоскостей их орбит к эклиптике от 0 до 90° были равновероятны. Требуется определить вероятность того, что при этом сумма наклонов десяти известных планет не считая Земли[14] должна быть заключена в заданных границах, например, между 0 и 90°. Принимая наклон планеты за вещь А, к которой относится формула (110.1), следует принять интервал $2g$ возможных значений А, равный 90°, а в самой формуле положить $\alpha = 0$, $\beta = 1$ и $\mu = 10$, так что $P = 1/10!$.

Эта дробь равна примерно четверти миллионной, и поэтому сумма наклонов, меньшая прямого угла, оказывается совершенно неправдоподобной; следует считать, что эта сумма должна была бы несомненно превышать 90°. Напротив, она доходит примерно лишь до 82°, а поскольку она претерпевает только очень небольшие периодические вариации, то гипотеза равной вероятности всех наклонов при образовании планет неприемлема. Нет никакого сомнения, что некая причина их образования должна была сделать наименьшие наклоны намного более вероятными.

Наклоны планет здесь считаются независимыми от любого направления движения планет, т. е. как у Земли вокруг Солнца, или противоположного. Если при образовании планет оба эти направления были равновероятны, то вероятность движения 10 планет, отличных от Земли, в её же направлении была бы равна 1/2 в десятой степени, т. е. ниже одной тысячной. Это также означает, что равный шанс обоих противоположных направлений очень маловероятен и доказывает, что неизвестная причина образования планет должна была привести к высокой вероятности движения всех планет в одном и том же направлении.

Примем теперь за А эксцентриситет планетной орбиты, и предположим, что первоначально все его значения от нуля до единицы были равновероятны. Тогда вероятность того, что сумма известных эксцентриситетов должна содержаться, к примеру, между нулём и 5/4, т. е. при $\alpha = 0$, $\beta = 1.25$ и $\mu = 11$, оказалась бы по формуле (110.1) равной $P = (1/11!)(1.25^{11} - 0.25^{11})$. Эта вероятность ниже трёх миллионных, и, напротив, исключительно вероятно, что сумма эксцентриситетов должна превысить 1.25. Тем не менее, эта сумма, которая претерпевает лишь небольшие периодические вариации, немного меньше 1.15. Гипотеза равной вероятности всех возможных значений А таким образом неприемлема, и вне сомнения, что некая причина образования планет сделала наименьшие эксцентриситеты, так же, как и наименьшие наклоны, намного более вероятными.

111. Сегодня насчитывают 138 комет, наблюдённых с 240-го года н. э. [до н. э.], для которых астрономы вычислили параболические элементы так тщательно, как было возможно. Из них 71 движется в прямом направлении и 67 – в попятном. Небольшое отличие этих чисел уже доказывает, что неизвестная причина образования комет не привела к их преимущественному движению в каком-то едином направлении. Сумма наклонов этих 138 комет к эклиптике доходит почти до 6752°, т. е.[15] превышает 75 прямых углов почти на 2°. Чтобы установить, должна ли она так мало отличаться от этой величины, будь все наклоны от нуля до 90° равновероятны, следует в формуле (110.1) принять α и β, немного уклоняющиеся от 75 в ту и другую сторону. Но это приведёт к невозможности численного вычисления. Чтобы при той же гипотезе определить вероятность P того, что сумма наклонов орбит всех наблюдённых комет должна содержаться в заданных границах, следует обратиться к формуле (101.3).

Я приму наклон кометной орбиты к эклиптике за вещь А. Границы возможных значений А, обычно обозначаемые через a и b, будут $a = 0$ и $b = 90°$, причём все эти значения считаются равновероятными. Формула (101.3) выразит вероятность P того, что среднее из большого числа μ наблюдённых наклонов в градусах окажется (ср. § 102) между $45 \mp 90u/\sqrt{6\mu}$. Приняв u = 1.92 и положив μ = 138, мы получим $P = 0.99338$ для вероятности, что при гипотезе равенства шансов всех возможных наклонов средний наклон 138 наблюдённых комет не выйдет за границы 45° ± 6°. Можно ставить почти 150 против одного, что это среднее должно оказаться между 39° и 51°, фактически же оно равно 48°55′. Таким образом, нельзя верить, что неизвестная причина образования комет придала их различным наклонам неравные вероятности.

Не вводя никаких гипотез о законе вероятностей этих наклонов, формула (101.3) также выражает вероятность того, что средний наклон большого числа μ комет, которые будут наблюдены впоследствии, отстоит от указанного значения 48°55′ лишь на число градусов, содержащихся в пределах (ср. § 107)

$$\mp \frac{ul\sqrt{138+\mu'}}{\sqrt{138\mu'}}.$$

По вычисленным наклонам 138 комет значение l, которое содержится в этих границах, племянник г-на Бювара[16] установил равным 34°49′. Полагая, к примеру, $\mu' = \mu$ и принимая как выше $u = 1.92$, можно будет ставить 150 против одного на то, что разность средних наклонов 138 наблюдённых и стольких же новых комет окажется в границах с интервалом ± 8°21′. Несомненно, что по сравнению с кометами, орбиты которых удалось вычислить, число μ' существующих комет исключительно велико, и предыдущие границы примут почти такой вид: $\mp ul/\sqrt{138}$. Они оказались более тесными в отношении 1: $\sqrt{2}$, чем при условии $\mu' = \mu$. Полагая неизменно $u = 1.92$, мы получим почти равную 150/151 вероятность того, что упомянутая выше разность будет заключена в границах ± 5°42′.

Если разделить поровну множество наблюдённых комет на две серии, состоящих из 69 более ранних и других 69 более современных, то в первой половине средний наклон будет равен 49°12′ и 48°38′ – во второй, с различием между ними чуть более полу-градуса. Этот пример весьма подходит, чтобы доказать, что средние значения одной и той же вещи соответствуют друг другу, даже если имеющиеся количества наблюдений не исключительно велики, а наблюдённые значения сильно отличаются друг от друга; здесь наименьший и наибольший наклоны равны 1°41′ и 89°48′. Средние наклоны 71 комет с прямым движением и 67 – с попятным движением больше отличаются друг от друга; первое среднее равно 47°3′, а второе – 50°54′.

Если в северном полушарии восставить перпендикуляр из центра Солнца к плоскости эклиптики, он пересечёт небесную сферу в северном полюсе эклиптики. И если в том же полушарии из этого центра восставить перпендикуляр к плоскости кометной орбиты, он пересечёт небесную сферу в северном полюсе этой орбиты. Угловое расстояние между этими полюсами будет наклоном указанной орбиты к эклиптике. Но нельзя смешивать, как это сделал уважаемый переводчик Гершеля (Herschel 1834), предположение, что все точки небесной сферы могут с одной и той же вероятностью быть полюсами кометных орбит с гипотезой равной вероятности всевозможных наклонов комет.

Пусть a и b будут круговыми зонами в северном полушарии одной и той же бесконечно малой ширины с общим центром в северном полюсе эклиптики, угловые расстояния которых от этого полюса равны α и β. Обозначим через p вероятность того, что случайная точка этого полушария принадлежит зоне a, и через q её принадлежность зоне b. Ясно, что дроби p и q

находятся в отношении a:b, а поэтому и в отношении синусов углов α и β. В соответствии с гипотезой равной возможности всем точкам небесной сферы быть полюсами кометных орбит, p и q выражают шансы расстояний α и β этих полюсов от эклиптики, или, иначе говоря, шансы кометных наклонов, равных этим расстояниям. Итак, в соответствии с принятой гипотезой, шансы различных наклонов пропорциональны синусам самих наклонов, но не равны друг другу. Шанс наклона в 90° оказывается вдвое большим, чем наклона в 30°, и оба бесконечно велики по сравнению с шансом бесконечно малого наклона[17].

112. Чтобы закончить эту главу, мы приводим сводку доказанных в ней и в предыдущей главе вероятностных формул. Число испытаний μ предполагается очень большим. Оно состоит из двух частей, m и n, которые также считаются очень большими. Формулы тем более точны, чем значительнее число μ, и они станут совершенно строгими при бесконечном μ.

112/1. Пусть p и q – постоянные на всём протяжении испытаний шансы противоположных событий E и F, так что $p + q = 1$. Обозначим через U вероятность того, что E появится m раз, а F – n раз при $\mu = m + n$. В соответствии с § 69

$$U = \left(\frac{\mu p}{m}\right)^m \left(\frac{\mu q}{n}\right)^n \sqrt{\frac{\mu}{2\pi mn}}. \qquad \text{(a)}$$

Эта формула была сведена в § 79 к

$$U = \frac{\exp(-v^2)}{\sqrt{2\pi\mu pq}},$$

притом

$$m = \mu p - v\sqrt{2\mu pq}, \; n = \mu q + v\sqrt{2\mu pq} \; .$$

Здесь v положительная или отрицательная величина, очень малая по сравнению с √μ. В этой форме формула равным образом пригодна при изменении шансов E и F от одного испытания к другому, если в соответствии с формулой (95.1) принять за p и q их средние значения по всей серии μ последовательных испытаний.

112/2. Пусть события E и F с неизвестными шансами p и q произошли m и n раз в μ испытаниях. Обозначим через U' вероятность, что они наступят m' и n' раз в будущих $\mu' = m' + n'$ испытаниях. Числа m' и n' пропорциональны m и n, так что

$$m' = \mu' m/\mu,\ n' = \mu' n/\mu.$$

При любом числе μ' будет иметь место формула (§ 71)

$$U' = \sqrt{\frac{\mu}{\mu+\mu'}}\ U_1, \qquad \text{(b)}$$

в которой U' – вероятность будущего события при условии, что m/μ и n/μ достоверно являются шансами E и F, т. е. что

$$U_1 = C_{\mu'}^{m'} (m/\mu)^{m'} (n/\mu)^{n'}.$$

112/3. Постоянные шансы p и q событий E и F заданы, а P – вероятность того, что в $\mu = m + n$ испытаниях E появится не менее m раз, а F – не более n раз. И (§ 77) при $q/p > n/(m + 1)$ и $q/p < n/(m + 1)$

$$P = \frac{1}{\sqrt{\pi}}\int_k^{\infty}\exp(-t^2)dt + \frac{(\mu+n)\sqrt{2}}{3\sqrt{\pi\mu mn}}\exp(-k^2)\ , \qquad (\mathrm{c}_1)$$

$$P = 1 - \frac{1}{\sqrt{\pi}}\int_k^{\infty}\exp(-t^2)dt + \frac{(\mu+n)\sqrt{2}}{3\sqrt{\pi\mu mn}}\exp(-k^2). \qquad (\mathrm{c}_2)$$

Здесь k – положительная величина, квадрат которой равен

$$k^2 = n\ln\frac{n}{q(\mu+1)} + (m+1)\ln\frac{m+1}{p(\mu+1)}.$$

112/4. Обозначим через R вероятность того, что количества появлений событий E и F в μ испытаниях не выйдут за границы

$$\mu p \mp u\sqrt{2\mu pq},\ \mu q \pm u\sqrt{2\mu pq}.$$

Здесь u – положительная величина, очень малая по сравнению с $\sqrt{\mu}$. Тогда (§ 79)[18]

$$R = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt + \frac{1}{\sqrt{2\pi\mu pq}}\exp(-u^2). \qquad \text{(d)}$$

Обратно, если шансы p и q неизвестны, а E и F произошли m и n раз в $\mu = m + n$ испытаниях, то (§ 83)

$$R = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt + \sqrt{\frac{\mu}{2\pi mn}}\exp(-u^2) \qquad \text{(e)}$$

окажется вероятностью того, что значения p и q не выйдут за границы

$$\frac{m}{\mu} \pm \frac{u}{\mu}\sqrt{\frac{2mn}{\mu}}, \quad \frac{n}{\mu} \mp \frac{u}{\mu}\sqrt{\frac{2mn}{\mu}}.$$

112/5. В двух сериях, состоящих из очень больших чисел μ и μ' испытаний, событие E произошло или произойдёт m и m' раз и n и n' раз наступило или наступит событие F. Обозначим через u положительную величину, очень малую по сравнению с $\sqrt{\mu}$ и $\sqrt{\mu'}$. Тогда вероятность того, что разность $m/\mu - m'/\mu'$ не выйдет за границы

$$\mp\frac{u\sqrt{2(\mu^3 m'n' + \mu'^3 mn)}}{\mu\mu'\sqrt{\mu\mu'}},$$

а $n/\mu - n'/\mu'$ не выйдет за те же границы с противоположным знаком, будет в соответствии с § 87 равна w:

$$w = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt +$$

$$\frac{\sqrt{\mu\mu'}}{\sqrt{2\pi m'n'(\mu + \mu')}}\exp\left[-\frac{u^2(\mu^3 m'n' + \mu'^3 mn)}{\mu^2 m'n'(\mu+\mu')}\right]. \qquad \text{(f)}$$

Поскольку $m/\mu \approx m'/\mu'$, а $n/\mu \approx n'/\mu'$, можно без ощутимого изменения значений w поменять местами величины μ', m', n' и μ, m, n в его последнем члене, который всегда будет небольшой дробью. Эта формула, если по крайней мере отвлечься от её последнего члена (§ 109), подходит к общему случаю, когда шансы E и F меняются от одного испытания к другому, лишь бы возможные известные или неизвестные причины этих событий не испытали никаких изменений в этих двух сериях, – если только существование этих причин сохраняет те же вероятности, а каждая из них неизменно придаёт один и тот же шанс появлению E, равно как и F.

112/6. События E и F неизменно появляются m и n раз в μ испытаниях. Пусть и два других противоположных событий E_1 и F_1 появляются m_1 и n_1 раз в μ_1 испытаниях, также очень многочисленных. Примем, что

$$m_1/\mu_1 - m/\mu = \delta,$$

где δ – очень малая положительная или отрицательная дробь и обозначим через p и p_1 неизвестные и постоянные шансы появления E и E_1 и через λ, – вероятность того, что p_1 превышает p по меньшей мере на малую и заданную положительную дробь ε. Далее, пусть u будет положительной величиной и положим, что

$$\frac{(\varepsilon-\delta)\mu\mu_1\sqrt{\mu\mu_1}}{\sqrt{2(\mu^3 m_1 n_1 + \mu_1^3 mn)}} = \pm\, u.$$

Знак здесь совпадает со знаком сомножителя ε – δ. В соответствии с § 88

$$\lambda = \frac{1}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt,\ \lambda = 1 - \frac{1}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt\,. \qquad \text{(g)}$$

Первое из этих выражений относится к случаю, при котором ε – δ положительно, второе из них – к отрицательной разности ε – δ. Они же также будут выражать вероятность того, что неизвестный шанс p наступления события E превосходит отношение m/μ, данное наблюдениями, на известную дробь w, для чего достаточно, чтобы

$$u = \pm\,(w - \frac{m}{\mu})\frac{\mu\sqrt{\mu}}{\sqrt{2mn}}.$$

Первую или вторую формулу следует принимать в зависимости от того, положительна или отрицательна разность $w - m/\mu$.

112/7. Пусть шансы противоположных событий E и F изменяются от одного испытания к другому. Значения этих шансов в испытании i обозначим через p_i и q_i, так что $p_i + q_i = 1$. Примем для сокращения записи, что

$$\sum p_i/\mu = p,\ \sum q_i/\mu = q,\ 2\sum p_i q_i/\mu = k^2,$$

где суммы распространяются от $i = 1$ до μ.

Далее, положим, что события E и F появились m и n раз в μ испытаниях и обозначим через u положительную и очень малую по сравнению с $\sqrt{\mu}$ величину. Тогда (§ 96)

$$R = 1 - \frac{2}{\sqrt{\pi}}\int_u^\infty \exp(-t^2)dt + \frac{\exp(-u^2)}{k\sqrt{\pi\mu}} \qquad \text{(h)}$$

выразит вероятность того, что отношения m/μ и n/μ не выйдут за границы $p \mp uk/\sqrt{\mu}$, $q \pm uk/\sqrt{\mu}$. Этот вывод совпадает с формулой (d) для частного случая постоянных шансов.

112/8. Пусть некоторая вещь A может принимать все значения в границах $h \pm g$, притом все они равно и единственно возможны. Обозначим через P вероятность того, что при некотором числе i испытаний сумма значений A оказалась заключённой также в заданных границах $c \pm \varepsilon$. Тогда (§ 99)

$$2(2g)^i P = \frac{\Gamma - \Gamma_1}{i!}, \qquad \text{(i)}$$

где

$$\Gamma = \pm(ih + ig - 4g - c + \varepsilon)^i \mp i(ih + ig - 2g - c + \varepsilon)^i \pm$$
$$C_i^2(ih + ig - 4g - c + \varepsilon)^i \mp C_i^3(ih + ig - 6g - c + \varepsilon)^i \pm \ldots$$

и Γ_1 равно Γ с изменённым знаком ε. В каждом члене этой формулы следует выбирать верхний знак, если величина, возводимая в степень, положительна и нижний знак в противном случае. Величины g и ε положительны, а h и c могут быть и положительны, и отрицательны.

112/9. Каков бы ни был закон вероятностей возможных значений вещи А при каждом испытании, и каким бы образом он не менялся бы от одного испытания к другому, обозначая через s сумму значений А в очень большом числе μ испытаний, мы имеем (§ 101)

$$P = 1 - \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} \exp(-t^2) dt \qquad \text{(k)}$$

для вероятности среднему s/μ значению А находиться в границах $k \mp 2u\sqrt{h/\mu}$. Здесь u обозначает положительную величину, очень малую по сравнению с $\sqrt{\mu}$, h положительно, а k и h зависят от вероятностей значений А во всей серии испытаний. Если указанные вероятности постоянны и одинаковы для всех возможных значений А между заданными границами a и b и равны нулю вне этих границ, то

$$k = (a + b)/2,\ h = (b - a)/2\sqrt{6}.$$

Если А принимает только конечное число v значений $c_1, c_2, \ldots, c_v$, и эти значения постоянны и равновероятны, то

$$k = \frac{1}{v}(c_1 + c_2 + \ldots + c_v)$$
$$h = \frac{1}{2v^2}[v(c_1^2 + c_2^2 + \ldots + c_v^2) - (c_1 + c_2 + \ldots + c_v)^2].$$

112/10. Пусть λ_n будет значением А в n-м испытании. Мы примем, что

$$\sum\lambda_n/\mu = \lambda,\ \sum(\lambda_n - \lambda)^2/\mu = l^2/2,$$

где суммы распространяются от $n = 1$ до μ. Предположим, что ни одна причина всех возможных значений А не испытала никаких изменений ни в своей вероятности, ни в шансах, которые она придает каждому из этих значений. Но тогда существует особая величина γ, к которой среднее s/μ значение А неограниченно приближается по мере всё большего возрастания μ, и которую оно достигнет, если μ станет бесконечным. И формула (k) выразит вероятность того, что γ заключена в границах (§ 106)

$$\frac{s}{\mu} \mp \frac{ul}{\sqrt{\mu}},$$

которые не содержат никаких неизвестных.

112/11. Пусть во второй серии из очень большого числа μ' испытаний s' будет суммой значений A, и l' – тем, чем станет l, которое относится к первой серии. Формула (k) также выразит вероятность того, что разность $s'/\mu' - s/\mu$ двух средних будет находиться в границах (§ 107)

$$\mp \frac{u\sqrt{l'^2\mu + l^2\mu'}}{\sqrt{\mu\mu'}}.$$

В случае примерного равенства l' и l эта же формула выразит вероятность того, что среднее для второй серии s'/μ' окажется в границах

$$\frac{s}{\mu} \mp \frac{ul\sqrt{\mu + \mu'}}{\sqrt{\mu\mu'}},$$

которые зависят лишь от результатов первой серии и заданной величины u и которые станут тем уже, чем больше μ' превосходит μ.

112/12. Пусть для установления значения одной и той же вещи A проделаны многие серии испытаний из очень больших чисел μ, μ', μ'', … Суммы значений A, полученных в этих сериях, обозначим через s, s', s'', … Предыдущая величина l будет, как и раньше, относиться к первой серии, в последующих же сериях соответствующие величины пусть будут l', l'', … Предположим, что причины погрешностей измерений изменяются от одной серии к другой, но что, по мере дальнейшего возрастания чисел μ, μ', μ'', …, все средние s/μ, s'/μ', s''/μ'', … тем не менее неограниченно стремятся к одной и той же неизвестной величине γ, которая является истинным значением A. Это имеет место, если в одной или нескольких сериях наблюдений указанные причины не приводят к неравной вероятности равных по величине и противоположных по знаку погрешностей[19]. Итак, формула (k) выразит также вероятность того, что A заключено в границы (§ 108)

$$\frac{sq}{\mu}+\frac{s'q'}{\mu'}+\frac{s''q''}{\mu''}+\ldots\mp\frac{u}{D}, \qquad (112.1)$$

где принято, что

$$\frac{\mu}{l^2}+\frac{\mu'}{l'^2}+\frac{\mu''}{l''^2}+\ldots=D^2,\ \ \frac{\mu}{D^2l^2}=q,\ \ \frac{\mu'}{D^2l'^2}=q',\ \ \frac{\mu''}{D^2l''^2}=q'',\ \ldots$$

Более того, сумма в формуле (112.1), т. е. сумма средних s/μ, s'/μ', s''/μ'', … умноженных соответственно на величины q, q', q'', … окажется приближённым значением γ, самым благоприятным из тех, которые могут быть выведены из сочетания всех серий наблюдений, т. е. значением этой неизвестной, границы ошибок которой $\mp u/D$ при заданном значении u или при той же степени вероятности будут самыми тесными из возможных.

112/13. Наконец, пусть причины появления события Е остаются без изменения в течение испытаний, как и было выражено формулой (f). Тогда отношение m/μ числа появлений Е к числу испытаний μ неограниченно стремится к особой величине r и станет в точности равным ей, если μ окажется бесконечным. И эта формула (f), если пренебречь её последним членом, или формула (k) окажется вероятностью того, что неизвестное значение r будет заключено в границах (§ 109)

$$\frac{m}{\mu}\mp\frac{u\sqrt{2m(\mu-m)}}{\mu\sqrt{\mu}}.$$

113. Чтобы завершить обзор формул, необходимо было бы добавить те, которые относятся к вероятностям одной или многих величин, выведенных из очень большого числа линейных уравнений, соответствующих результатам такого же числа наблюдений. Но относительно этих остальных формул я отсылаю читателей к Лапласу (1812). Исходя из системы 126 уравнений, составленных Бюваром и относящихся к движению Сатурна по долготе[20], и прилагая к ним метод *наименьших квадратов*, он заключил, что можно ставить миллион против одного за то, что масса Юпитера, если принять массу Солнца за единицу, не отличается ни в одну сторону более, чем на сотую долю дроби 1/1070, см Laplace (1812/1886, pp. 516 – 519). Тем не менее, последующие наблюдения другой природы привели к величине, почти равной 1/1050, что превышает дробь 1/1070 примерно на 1/50 её значения и, как казалось, установили дефект исчисления вероятностей. По поводу массы в 1/1050 не может оставаться никаких сомнений, как заключил Энке по возмущениям кометы с периодом возвращения 1204 дней; как заключили Гаусс и Николаи[21] по возмущениям Весты и Юноны, и Эри по недавно измеренным элонгациям спутников Юпитера.

И всё-таки, если даже вычисления Лапласа установили массу этой планеты с вероятностью, очень близкой к достоверности, с недостатком в 1/50 относительно её истинной величины, то нельзя заключить, что сила притяжения Юпитера могла действовать на Сатурн меньше, чем на свои собственные спутники, на кометы и малые планеты. И этот результат также не произошёл от какой-либо неточности в использованных Лапласом формулах вероятности. Можно считать, что установление им чуть меньшей массы Юпитера оказалось результатом нескольких ошибочных членов в выражении столь сложных возмущений Юпитера, которые были уже несколько исправлены, притом что остальные члены могут потребовать других поправок. Важной целью *Небесной механики*, которую наверняка достигнет нынешняя работа Бювара[22], является полная переделка уже столь точных таблиц движений Сатурна и Юпитера.

Примечания

1. Сумма $\sum U(m, n)u^m v^n$ называется двойной производящей функцией u и v. См. также § 20.

2. Формула (95.1) выражает локальную предельную теорему для схемы испытаний с переменной вероятностью. За ней последовала интегральная предельная теорема, формулы (97.1) и (101.1). Здесь же заметим, что Пуассон исследовал и центральную предельную теорему, которую (если не считать её частного случая, теоремы Муавра – Лапласа) строго доказали только А. А. Марков и А. М. Ляпунов. См. §§ 102 – 103, а также другие его сочинения (1824, §§ 1 – 7; 1837). На основании указанных параграфов Пуассон доказал *второе основное предложение* закона больших чисел из § 52, но он явно недостаточно изложил условия теоремы, а с методической точки зрения изложение Пуассона малоудовлетворительно.

Hald (1998, pp. 317 – 327) подробно рассмотрел доказательство центральной предельной теоремы у Пуассона. Без обоснования он (с. 327) указал, что условия теоремы не были явно указаны в соответствии с *преобладающим обычаем*. Странно также, что он так и не сообщил, в чём доказательство Пуассона было недостаточно строгим. Строгости добились здесь, как известно, только Марков и Ляпунов.

3. Пуассон выводит формулу обращения для сумм разнораспределённых решётчатых случайных величин. Чуть раньше он применил характеристическую функцию.

4. По поводу исследования этого особого случая я отсылаю читателей к своему мемуару (1824), уже упомянутому в § 60. Автор

5. Эти предыдущие выражения см. в § 94.

6. В приводимой формуле не приведены подобные члены.

7. См. Прим. 6.

8. Пуассон далее вводит функцию Дирака. Впервые он (1811/1833, с. 637) применил её ещё раньше.

9 Эту фразу следовало включить чуть выше.

10. В Прим. 3 к гл. 3 мы указали, что Пуассон редко применял понятия дисперсии.

11. В этом параграфе Пуассон исследовал допустимую разность средних значений случайной величины в двух сериях испытаний с переменной вероятностью; в § 109 изучил разность средних частостей появления случайной величины. Разность вероятностей Пуассон (§ 88 гл. 3) рассмотрел только для серий испытаний по схеме Бернулли.

12. См. Прим. 18 к гл. 2.

13. Могу также упомянуть вероятность стрельбы в мишень, которую я рассмотрел в мемуаре, написанном до этой книги. Он выйдет […] (1837). Автор

14. Лаплас также исходил из данных по 10 планетам, а перечислил их Курно (1843/1970, с. 235), который включил в их число 4 малые планеты.

15. Автор не пояснил, почему сумма наклонов комет должна была приближённо равняться 75 прямым углам. Выше, он не указал, что эксцентриситеты планетных орбит обусловлены скоростью обращения планет вокруг Солнца и что возможно правильнее было бы рассуждать об этих скоростях.

16. А. Бювар (1767 – 1843), французский астроном. Но его племянника следовало бы назвать по имени.

17. В приведенном примечании автор качественно объясняет происхождение *аэролитов* (метеоритов) и падающих звёзд существованием *неисчерпаемого* множества очень малых тел, вращающихся вокруг Солнца, планет *или даже спутников*. В следующем параграфе автор приводит сводку важнейших формул, но не всегда в прежнем виде.

18. Автор не сравнил чётко точности прямой и обратной задач, см. формулы (d) и (e). Лишь Бейес доказал, что обратная задача менее точна (Шейнин 2011) Ни Якоб Бернулли, ни Муавр этого не заметили. См. также конец § 71 в гл. 3.

19. См. Прим. 18 к гл. 2.

20. Лаплас не смог бы определить массу Юпитера по наблюдениям Сатурна. На самом деле он (Приложение № 1 (1816) к его *Аналитической теории вероятностей*, см. Laplace (1812/1886, с. 516 – 519), исходил из 126 уравнений, относящихся к Юпитеру и 129 – относящихся к Сатурну. Он заключил, что в последующее столетие, если наблюдения будут обрабатываться тем же способом, его оценка массы Юпитера не изменится даже на 1/100. Как впоследствии Пуассон, он не учитывал неизбежного существования систематических ошибок. На фундаментальные результаты Гаусса в теории ошибок Лаплас почти никогда, а Пуассон вообще никогда, не ссылались.

21. F. B. G. Nicolai (1793 – 1846), немецкий астроном.

22. Небесная механика здесь упомянута как научная дисциплина, а не как сочинение Лапласа, но почему-то эти два слова были напечатаны курсивом.

Библиография

Курно О. (1843, франц.), *Основы теории шансов и вероятностей.* М. Перевод Н. С. Четверикова.

Шейнин О. Б. (2011), Обратный закон больших чисел. *Историко-математич. исследования*, вып. 14 (49), с. 212 – 219.

Hald A. (1998), *History of Mathematical Statistics from 1750 to 1930*. New York.

Herschel J. (1834), *Traité d'astronomie.* Перевод О. Курно с английского *Outlines of Astronomy*. London, 1849, 2nd edition.

Laplace P. S. (1816), Первое приложение к *Théorie analytique des probabilités*, 1812. *Oeuvr. Compl.*, t. 7. Paris, 1886, pp. 497 – 530.

Poisson S. D. (1811), *Traité de mécanique*, t. 1. Paris, 1833.

--- (1824), Sur la probabilité des résultats moyens d'observations. *Conn. des temps* pour 1827, pp. 273 – 302.

--- (1829), Sur la probabilité …Ibidem, pour 1832, pp. 3 – 22.

--- (1837), Sur la probabilité du tir a la cible. *Mémorial de l'artill.*, No. 4, pp. 59 – 94.

Гл. 5. Приложение общих правил вероятностей к решениям судов присяжных и суждениям трибуналов[1]

Опечатки/ошибки, не замеченные автором

1. В § 116, с. 327 оригинала. В последней скобке знаменателя должно быть $(1 - u')$, а не $(1 - u'')$.

2. В § 127, с. 350 оригинала. Левая часть второй формулы (14) должна быть Z_i, а не Z.

3. В § 130, с. 358 оригинала. Первый из интегралов от $U_i \varphi u du$: пределы второго интеграла в правой части равенства должны быть $-\infty$ и 0, а не ∞ и 0.

4. В § 131, с. 361 оригинала, строка 2 снизу. Подынтегральная функция в левой части формулы должна содержать U_i, а не U'_i.

5. В § 134, с. 370 оригинала. Разность отношений должна быть $a'_i/\mu' - a_i/\mu$, но вместо a_i было оставлено пустое место.

6. В § 135, с. 374 оригинала. После двух выключенных из общего текста значений a_4/μ в самом тексте a_4/μ ошибочно выписано в виде a^4/μ. В переводе исправлено. Также исправлены несколько таких же ошибок в §§ 140 и 143.

7. В § 139, с. 383 оригинала. Предпоследний член второй квадратной скобки формулы (139.2) должен быть $495t^4$, а не $495t^5$.

8. В § 149, с. 409 оригинала. Последний член первой квадратной скобки выражения $C\,35v^4(1 - v)$ должен быть равен $35v^4(1 - v)^3$.

В переводе все эти ошибки/опечатки исправлены.

114. В столь затруднительной проблеме уместно прежде всего рассмотреть простейшие случаи, а затем уже исследовать её во всей общности. И я предположу, что судит один только присяжный. Пусть k – вероятность априорной вины подсудимого. Эта вероятность устанавливается по предварительным сведениям и последующему обвинению. Я также обозначу через u вероятность того, что присяжный не ошибается при своём решении и через γ – вероятность того, что подсудимый будет осуждён. Это произойдёт, если он виновен и присяжный не ошибся, или невиновен, а присяжный ошибся.

По правилу § 5 вероятность первого случая равна ku, а второго – $(1 - k)(1 - u)$. Поэтому, в соответствии с правилом § 10, полная вероятность осуждения подсудимого равна

$$\gamma = ku + (1 - k)(1 - u), \qquad (114.1)$$

а оправдания – $(1 - \gamma)$. Это произойдёт, если подсудимый виновен, а присяжный ошибся, или невиновен и присяжный не ошибся. Поэтому

$$1 - \gamma = k(1 - u) + u(1 - k),$$

что можно было вывести из предыдущего. Вычитая второе равенство из первого, мы получим

$$2\gamma - 1 = (2k - 1)(2u - 1),$$

что доказывает, что $2\gamma - 1 = 0$ когда либо $(2k - 1)$, либо $(2u - 1)$ равно нулю. Эта разность будет положительна или отрицательна в соответствии с тем, окажутся ли знаки $(2k - 1)$ и $(2u - 1)$ одними и теми же или противоположными. Мы также имеем

$$\gamma = 1/2 + (2k - 1)(2u - 1)/2,$$

так что γ отличается от 1/2 на половину произведения $(2k - 1)(2u - 1)$, положительного или отрицательного.

После решения присяжного можно сформулировать две единственно возможные гипотезы: был ли подсудимый виновен или нет. Их вероятности, как и во всех предположительных случаях, определяются по правилу § 34, а их сумма равна единице, так что достаточно определить одну из них.

Пусть подсудимый осуждён, а вероятность первой гипотезы или его виновности равна p. По указанному правилу

$$p = \frac{ku}{ku + (1 - k)(1 - u)}, \qquad (114.2)$$

потому что наблюдённым событием является осуждение подсудимого, вероятность чего в соответствии с обеими гипотезами, как только что было указано, равна ku или $(1 - k)(1 - u)$.

При оправдании подсудимого обозначим через q вероятность второй гипотезы, т. е. его невиновности. Наблюдённым событием будет оправдание подсудимого, вероятность чего в соответствии с этой гипотезой равна $(1 - k)u$ или $k(1 - u)$ при противоположном предположении, и поэтому

$$q = \frac{(1-k)u}{(1-k)u + k(1-u)}. \qquad (114.3)$$

Знаменателями выражений p и q являются значения γ и $1 - \gamma$, так что

$$p = ku/\gamma, \; q = (1 - k)u/(1 - \gamma).$$

Следовательно, вероятность того, что присяжный не ошибся, равна

$$u = p\gamma + q(1 - \gamma),$$

что легко проверить. По существу это происходит в двух различных случаях: подсудимый будет осуждён и притом виновен, или будет оправдан и притом невиновен.

В соответствии с правилом § 9 о вероятности события, составленного из двух простых событий, шансы которых влияют друг на друга, вероятность первого случая равна γp, а второго – $(1 - \gamma)q$, однако (§ 10) полное значение u является суммой этих двух произведений. После того, как решение присяжного объявлено, вероятность того, что он не ошибся, равна p, если подсудимый осуждён, или q, если он оправдан. И если $k \neq 1/2$, она, как прежде, может равняться u только если $u = 0$ или 1.

Приведенные формулы предоставляют полное решение проблемы в случае только одного присяжного. Впрочем, она совпадает с проблемой вероятности события, удостоверенного одним свидетелем (§ 36). Событием, которое может быть истинным или ложным, является здесь виновность подсудимого. До объявления решения присяжного существовало некоторое основание считать, что это событие истинно: вероятность этого равнялась k, а $(1 - k)$ было вероятностью невиновности подсудимого. После решения присяжного появилось новое сведение, и k стало равным другой вероятности, p, если подсудимый признан виновным, или же $(1 - k)$ становится равным q в противном случае.

В каждом случае ясно, что предварительные вероятности k и $(1 - k)$ должны были повыситься, если шанс безошибочности присяжного превышает половину, и понизиться в противном случае. Иначе говоря, повыситься или понизиться в зависимости от того, будет ли $u > 1/2$ или $< 1/2$. Это следует из выражений для p и q, из которых можно вывести, что

$$p = k + \frac{k(1-k)(2u-1)}{\gamma},\ q = 1 - k - \frac{k(1-k)(2u-1)}{1-\gamma},$$

так что $p > k$, $p < k$, $q < (1-k)$, $q > (1-k)$ в соответствии с неравенствами $u > 1/2$ и $u < 1/2$. При $u = 1/2$ предварительные вероятности k и $(1 - k)$ нисколько не меняются.

Из последних выражений для p и q следует, что

$$p\gamma + q(1 - \gamma) = u = k\gamma + (1 - k)(1 - \gamma).$$

Если заранее каким-то образом, кроме вероятности k вины, известен шанс γ осуждения, то эти формулы могут служить для вычисления вероятности, что присяжный не ошибается. И они могут быть проверены при наблюдении, что присяжный не ошибается, если подсудимый виновен и осуждён, или если он невиновен и оправдан. Вероятности этих двух случаев до решения присяжного равны произведениям $k\gamma$ и $(1 - k)(1 - \gamma)$, и их сумма образует полное значение u.

Если $k = 1/2$, первые значения p и q сразу же становятся равными $p = q = u$. По существу, поскольку заранее нет никаких причин полагать подсудимого скорее виновным, чем невиновным, после решения присяжного наше основание верить в одно или другое не может отличаться от вероятности того, что он не ошибается. Если $k = 1$, т. е. если вероятность вины заранее считается достоверной, то $p = 1$ и $q = 0$. Каковы бы ни были это решение и шанс его безошибочности, виновность подсудимого окажется ещё более достоверной после решения. То же имеет место относительно невиновности подсудимого при $k = 0$, т. е. при её априорной достоверности. Но в этих случаях неизвестно, будет ли подсудимый осуждён или оправдан. В первом случае шанс осуждения $\gamma = u$ и $1 - u$ во втором. Эти шансы будут равны, как это и должно быть, вероятности того, что при $k = 1$ и 0 присяжный соответственно не ошибается и ошибается.

115. Предположим, что после решения присяжного подсудимый судится вторым присяжным, вероятность которого не ошибаться равна u'. Требуется определить вероятности c, b, a того, что подсудимый будет осуждён обоими; оправдан одним и осуждён другим; оправдан обоими.

Пусть γ' будет вероятностью осуждения подсудимого вторым присяжным после осуждения первым. Шанс первого осуждения равен γ, и потому вероятность осуждения обоими окажется равной $c = \gamma\gamma'$. Но при появлении перед вторым присяжным уже существует возникшая после решения первого вероятность p того, что подсудимый виновен. Поэтому значение γ' выводится из формулы (114.1) при подстановке p и u' вместо k и u. Таким образом,

$$\gamma' = pu' + (1 - p)(1 - u')$$

и из формул (114.1) и (114.2) будет следовать, что

$$c = kuu' + (1 - k)(1 - u)(1 - u').$$

Аналогичное рассуждение приводит к

$$a = k(1 - u)(1 - u') + (1 - k)uu'.$$

Объединяя эти две формулы, мы получаем вероятность того, что решения обоих присяжных то ли осудить, то ли оправдать, совпадают:

$$a + c = uu' + (1 - u)(1 - u').$$

Заметим, что эта полная вероятность не зависит от предварительной вероятности вины подсудимого.

Если подсудимый был оправдан первым присяжным, а вероятность его осуждения вторым обозначена через γ_1, то произведение $(1 - \gamma)\gamma_1$ выразит вероятность того, что эти противоположные решения произойдут в указанном порядке. Кроме того, $(1 - q)$ будет вероятностью того, что при появлении перед вторым присяжным после оправдания первым подсудимый виновен. Значение γ_1 определяется по формуле (114.1) при замене k и u на $(1 - q)$ и u', так что оказывается, что

$$\gamma_1 = (1 - q)u' + q(1 - u').$$

Учитывая значения $(1 - \gamma)$ и q, заданные формулами (114.1) и (114.3), мы имеем

$$(1 - \gamma)\gamma_1 = k(1 - u)u' + (1 - k)(1 - u')u.$$

Ясно, что, переставляя буквы u и u' в этом выражении, мы получим вероятность того, что решения присяжных противоположны, но были приняты в порядке, обратном только что предположенному. Добавление этой вероятности к предыдущей приводит к полной вероятности противоположных решений присяжных, принятых в каком-то порядке

$$b = (1 - u)\, u' + (1 - u')u.$$

Видно, что она, как и в случае совпадающих решений, не зависит от k. При $u = u' = 1/2$ обе эти вероятности тоже равны 1/2, и их сумма, $a + b + c$, неизменно равна единице, как это и должно было быть.

Вероятность вины подсудимого, осуждённого обоими присяжными, выражается формулой (114.2) при замене k и u на p и u', а вероятность его невиновности после оправдания обоими выводится из формулы (114.3) при замене k и u на $(1 - q)$ и u'. Обозначая эти две вероятности через p' и q', мы имеем

$$p' = \frac{pu'}{pu' + (1-p)(1-u')},\ q' = \frac{qu'}{qu' + (1-q)(1-u')}.$$

С учётом значений p и q, данных теми же самыми формулами (114.2) и (114.3), эти p' и q' становятся равными

$$p' = \frac{kuu'}{kuu' + (1-k)(1-u)(1-u')},\ q' = \frac{(1-k)uu'}{(1-k)uu' + k(1-u)(1-u')}.$$

Пусть также p_1 будет вероятностью вины подсудимого после его оправдания первым присяжным и осуждения вторым, а q_1 – вероятностью его невиновности после осуждения первым присяжным и оправдания вторым. Значение вероятности p_1 того, что подсудимый, оправданный первым присяжным, не является невиновным, выводится из формулы (114.2) при замене u и k на u' и $(1 - q)$. Значение q_1 выводится из формулы (114.3) при замене u и k на u' и p. Таким образом,

$$p_1 = \frac{(1-q)u'}{(1-q)u' + q(1-u')},\ q_1 = \frac{(1-p)u'}{(1-p)u' + p(1-u')}$$

и, с учётом этих же формул (114.2) и (114.3),

$$p_1 = \frac{k(1-u)u'}{k(1-u)u' + (1-k)(1-u')u},\ q_1 = \frac{(1-k)(1-u)u'}{(1-k)(1-u)u' + k(1-u')u}.$$

Вероятность того, что подсудимый, осуждённый первым присяжным и оправданный вторым, виновен, будет равна $(1 - q_1)$. Кроме того, ясно, что она должна будет выводиться из p_1 перестановкой u и u', что по существу и имеет место: в этом случае вероятность невиновности подсудимого, оправданного первым и осуждённым вторым, или $(1 - p_1)$, выводится из q_1 при перестановке u и u'.

При $u' = u$ будет $p_1 = k$ и $q_1 = (1 - k)$, что и должно было быть, потому что противоположные решения присяжных, обладающих одним и тем же шансом не ошибаться, не могут ничего изменить в основаниях веры в виновность или невиновность подсудимого, существовавших до этих решений.

116. Можно без труда обобщить эти рассуждения на последовательные решения какого-либо числа присяжных, обладающих известным шансом не ошибаться. Но к соответствующим результатам можно придти более просто следующим образом.

Для определённости я предположу, что присяжных трое. Обозначим их вероятности не ошибаться через u, u' и u'' и, как и раньше, через k – предварительную вероятность того, что подсудимый виновен. Для единогласного осуждения требуется, чтобы он действительно был виновен, и никто из присяжных не ошибся, либо, чтобы он был невиновен, а все трое присяжных ошиблись. Полная вероятность такого осуждения поэтому будет равна

$$kuu'u'' + k(1 - u)(1 - u')(1 - u'').$$

Равным образом видно, что вероятность единогласного оправдания будет равна

$$k(1 - u)(1 - u')(1 - u'') + (1 - k)uu'u''.$$

Сумма этих двух выражений окажется вероятностью единогласного решения то ли осуждения, то ли оправдания:

$$uu'u'' + (1 - u)(1 - u')(1 - u'').$$

Она не зависит от k, и это равным образом относится к любому числу присяжных.

Подсудимый может быть осуждён двумя присяжными и оправдан третьим в трёх различных случаях в зависимости от того, будет ли шанс не ошибаться у этого третьего равен u, u' или u''. Он также может быть оправдан двумя и осуждён третьим, снова в тех же трёх случаях. Нетрудно видеть, что вероятности всех шести комбинаций равны

$$ku'u''(1 - u) + (1 - k)(1 - u')(1 - u'')u,$$
$$kuu''(1 - u') + (1 - k)(1 - u)(1 - u'')\, u',$$
$$kuu'(1 - u'') + (1 - k)(1 - u)(1 - u')u'',$$
$$k(1 - u')(1 - u'')u + (1 - k)(1 - u)u'u'',$$
$$k(1 - u)(1 - u'')u' + (1 - k)(1 - u')uu'',$$
$$k(1 - u)(1 - u')u'' + (1 - k)(1 - u'')uu'.$$

Суммируя эти шесть величин, мы получим полную вероятность того, что решения не были единогласными:

$$u'u''(1 - u) + uu''(1 - u') + uu'(1 - u'') + (1 - u')(1 - u'')u + (1 - u)(1 - u'')u' + (1 - u)(1 - u')u''.$$

Видно, что она не зависит от k. Сумма полных вероятностей единогласных и не единогласных решений должна равняться единице, и их выражения удовлетворяют этому условию.

После вынесения решений можно легко вывести вероятность вины подсудимого, которая, вообще говоря, будет отличаться от существовавшей прежде. Если, к примеру, подсудимый осуждён двумя присяжными, шансы не ошибаться у которых равнялись u и u' и оправдан третьим, вероятность этого результата будет равна $kuu'(1 - u'')$ или $(1 - k)(1 - u)(1 - u')u''$ в соответствии с гипотезами его вины и невиновности. Поэтому, по правилу § 34, вероятность его вины оказывается равной

$$\frac{kuu'(1-u'')}{kuu'(1-u'')+(1-k)(1-u)(1-u')u''}.$$

При $u' = u''$ она становится независимой от общего значения u' и u''. То же будет иметь место, если подсудимый осуждается только одним присяжным, шанс не ошибаться которого равен u. И по существу после объявленного решения различные варианты решения двумя другими присяжных уже не могут повлиять на основания моей веры в виновность или невиновность подсудимого. Действительно, нет большей причины для повышения, а не понижения вероятности его вины, потому что шансы не ошибаться у двух последних присяжных равны по предположению.

Приведенные формулы равно применимы к случаю, при котором присяжные объединены и судят после [совместного] обдумывания, а не последовательно и без взаимного общения. Обсуждение может прояснить для них положение и, вообще говоря, повысить их вероятность не ошибаться[2]. Значения u, u' и u'', которые относятся к этим двум случаям, могут не быть теми же самыми, и во втором из них могут менее уклоняться от единицы, чем в первом.

117. Рассмотрим, в частности, случай, при котором шансы не ошибаться одни и те же у всех присяжных. К нему мы затем приведём общий случай определения вероятности числа осуждений при очень большом числе решений.

Обозначим через u эту заданную вероятность присяжных не ошибаться, через n их число; и, наконец, через k – вероятность вины подсудимого, существующую до их решения. Пусть также $i = 0, 1, 2, \ldots, n$ и γ_i – вероятность осуждения подсудимого $(n - i)$ присяжными и его оправдания i присяжными.

Для наступления указанного составного события необходимо, чтобы подсудимый был виновен, $(n - i)$ присяжных не ошиблись, а i из них ошиблись; или же, чтобы подсудимый не был виновен, $(n - i)$ ошиблись, а i из них не ошиблись. Вероятность первого случая равна произведению $ku^{n-i}(1 - u)^i$, умноженному на число раз, которым можно выбрать i присяжных из их общего числа n. Вероятность второго случая равна произведению $(1- k)u^i(1 - u)^{n-i}$, умноженному на число раз, которым можно выбрать $(n - i)$ присяжных из их общего числа. Это число совпадает с соответствующим числом в первом случае […] и в результате

$$\gamma_i = N_i[ku^{n-i}(1 - u)^i + (1 - k)u^i(1 - u)^{n-i}] \qquad (117.1)$$

[через N_i автор обозначил число сочетаний из n по i] окажется полным значением γ_i.

Если $n - i > i$, то пусть $n - 2i = m$, так что подсудимый осуждается большинством в m голосов. Если же i присяжных осуждают его, а $(n - i)$ оправдывают, то он оправдывается тем же большинством. Вероятность такого оправдания, которую я обозначаю через δ_i, определяется по значению γ_i при перестановке чисел $(n - i)$ и i, что не меняет значения N_i. Таким образом

$$\delta_i = N_i[ku^i(1 - u)^{n-i} + (1 - k)u^{n-i}(1 - u)^i]. \qquad (117.2)$$

Складывая два последних уравнения, мы получаем сумму

$$\gamma_i + \delta_i = N_i[u^{n-i}(1 - u)^i + u^i(1 - u)^{n-i}],$$

которая не зависит от k. Вероятность либо осуждения, либо оправдания заданным большинством m голосов, таким образом не зависит от заранее предполагаемой вины подсудимого. В частном случае $u = 1/2$ взятые по отдельности вероятности γ_i и δ_i также независимы от k, а их совпадающее значение равно

$$\gamma_i = \delta_i = N_i/2^n.$$

При $k = 1/2$ они также равны друг другу при любом значении u.

118. Пусть теперь c_i будет вероятностью осуждения подсудимого не менее, чем $(n - i)$ присяжными и оправдания не более, чем i из них, т. е. осуждёния большинством по меньшей мере в m голосов. Обозначим также через d_i вероятность того, что подсудимый оправдан не менее, чем $(n - i)$ голосами против не более, чем i голосов. По правилу § 10 мы имеем

$$c_i = \gamma_0 + \gamma_1 + \gamma_2 + \ldots + \gamma_i, \; d_i = \delta_0 + \delta_1 + \delta_2 + \ldots + \delta_i,$$

а ввиду предыдущих формул оказывается, что

$$c_i = kU_i + (1 - k)U_i, \; d_i = kV_i + (1 - k)U_i. \qquad (118.1)$$

Примем для сокращения письма

$$N_0u^n + N_1u^{n-1}(1 - u) + N_2u^{n-2}(1 - u)^2 + \ldots + N_iu^{n-i}(1 - u)^i = U_i,$$
$$N_0(1 - u)^n + N_1(1 - u)^{n-1}u + N_2(1 - u)^{n-2}u^2 + \ldots + N_i(1 - u)^{n-i}u^i = V_i,$$

где U_i – заданная функция u, причём при замене u на $(1 - u)$ эта функция переходит в V_i. В то же время

$$c_i + d_i = U_i + V_i$$

оказывается не зависящей от k вероятностью того, что подсудимый будет либо осуждён, либо оправдан большинством по меньшей мере m голосов.

Если в выражении для d_i заменить i на $(n - i - 1)$, то окажется, что

$$U_i + V_{n-i-1} = 1,\ c_i + d_{n-i-1} = 1.$$

Пусть для осуждения требуется по меньшей мере $(n - i)$ голосов, тогда подсудимый будет оправдан, если так решат не более, чем $(n - i - 1)$ присяжных. Это означает, что одно из двух событий, вероятности которых равны c_i и d_{n-i-1} наверняка произойдёт.

При нечётном n и, следовательно, при $n = 2i + 1$ и $m = 1$ будет

$$U_i + V_i = [u + (1 - u)]^n = 1,\ c_i + d_i = 1.$$

Это означает, что подсудимый будет наверняка либо осуждён, либо оправдан большинством по меньшей мере в один голос, что и так очевидно. Если n чётно, наименьшее большинство окажется равным $m = 2$, что соответствует равенству $n = 2i + 2$. Тогда

$$U_i + V_i = [u + (1 - u)]^n - N_{i+1}u^{i+1}(1 - u)^{i+1}$$

$$c_i + d_i = 1 - \frac{(2i+2)(2i+1)2i...(i+2)}{(i+1)!}[u(1-u)]^{i+1}.$$

Всё же осуждение или оправдание большинством по меньшей мере в два голоса не является достоверным, что очевидно ввиду возможного равного разделения голосов за оправдание и осуждение. Вероятность этого исключительного случая определяется вычитанием предыдущего значения $c_i + d_i$ из единицы, и она независима от k. Обозначив её через H_i, мы можем записать её в виде

$$H_i = \frac{(2i+2)![u(1-u)^{i+1}]}{[(i+1)!]^2}.$$

Функция $u(1 - u)$ достигает максимума, равного 1/4, при $u = 1/2$. Вероятность H_i убывает по мере дальнейшего отклонения u от 1/2. Она также монотонно убывает по мере возрастания i. Действительно, это выводится из выражения

$$H_{i+1} = \frac{(2i+3)(2i+4)u(1-u)}{(i+1)^2} H_i .$$

После прохождения указанного выше максимума отношение $H_{i+1}:H_i$ будет неизменно меньше единицы; наибольшее значение H_i соответствует случаю $u = 1/2$ и $i = 0$ и равно 1/2.

Если $i + 1$ – очень большое число, то (§ 67) […] и

$$H_i = \frac{[4u(1-u)]^{i+1}}{\sqrt{\pi(i+1)}}[1 - \frac{1}{8(i+1)} + ...]$$

будет приближённым значением этой величины, и, как заметно, окажется очень малой дробью, если u ощутимо отклоняется от 1/2 или $4u(1 - u)$ – от единицы. При $u = 1/2$ и, к примеру, $i + 1 = 6$ или $n = 12$ эта формула, если ограничиться первыми двумя членами в квадратных скобках, приведёт к значению 230.94 …/1024. Это очень немного отличается от точного значения 231/1024, хоть $i + 1$ и не является весьма большим числом.

Сумма $U_i + V_i = G_i$ принимает значения, не превышающие единицу, так что разность $1 - G_i$ будет положительна или равна нулю. А поскольку выражение для c_i можно записать в виде

$$c_i = k - k(1 - G_i) - (2k - 1)V_i,$$

то при $2k - 1 > 0$ или $k > 1/2$ будет $c_i < k$. Поэтому в обычных случаях, при которых до решения присяжных имеются основания верить скорее в вину, а не в невиновность подсудимого, шанс его осуждения большинством по меньшей мере в один голос, т. е. любым большинством, всегда окажется ниже этой предварительной вероятности его вины. Предположим, например, что заранее можно было ставить 4 против одного за его вину, но вот соотношение ставок за его осуждение окажется меньшим.

Ясно, что это предложение не зависит ни от шансов ошибок присяжных, ни от значений u, отличных от единицы. При $u = 1$ при любом i будет $U_i = 1$, $V_i = 0$, $c_i = k$, $d_i = 1 - k$, а при $u = 0$ $U_i = 0$,

$V_i = 1$, $c_i = (1 - k)$, $d_i = k$. При этих двух крайних значениях u осуждение или оправдание очевидно не может произойти кроме как единогласно. Это следует из формул (117.1) и (117.2), которые, исключая случай $i = 0$, приведут к $\gamma_i = \delta_i = 0$.

119. Сохраняя все предыдущие обозначения, дополнительно представим через p_i вероятность того, что подсудимый виновен, если он осуждён $(n - i)$ присяжными против решения i из них, т. е. большинством в m голосов, и через q_i – вероятность, что невиновен, если оправдан тем же большинством. Иначе говоря, p_i и q_i оказываются вероятностями того, что решение осудить или оправдать подсудимого большинством m голосов при n присяжных является верным. В первом случае вероятность наблюдённого события, т. е. осуждения, равна $N_i k u^{n-i}(1 - u)^i$ или $N_i(1 - k)^i(1 - u)^{n-i}u^i$ в соответствии с тем, виновен ли подсудимый или нет. По правилу § 34, при сокращении множителя N_i в числителе и знаменателе,

$$p_i = \frac{ku^{n-i}(1-u)^i}{ku^{n-i}(1-u)^i + (1-k)(1-u)^{n-i}u^i}. \qquad (119.1)$$

При оправдании

$$q_i = \frac{(1-k)u^{n-i}(1-u)^i}{(1-k)u^{n-i}(1-u)^i + k(1-u)^{n-i}u^i}. \qquad (119.2)$$

Если $k = 1/2$, то $p_i = q_i$. Действительно, если заранее нет оснований полагать подсудимого скорее виновным, чем невиновным, ясно, что верность решения дела тем же самым большинством обладает одной и той же вероятностью и при осуждении, и при оправдании. При $u = 1/2$ будет $u^{n-i}(1 - u)^i = (1 - u)^{n-i}u^i$ и, как и должно было быть, $p_i = k$ и $q_i = 1 - k$ каковы бы ни были n и i.

Пусть в формулах (119.1) и (119.2) $u = t/(1 + t)$, $1 - u = 1/(1 + t)$. Замечая, что $n = m + 2i$, мы имеем

$$p_i = \frac{kt^m}{kt^m + 1 - k},\ q_i = \frac{(1-k)t^m}{(1-k)t^m + k},$$

что доказывает, что вероятность верности решения при прочих равных условиях зависит лишь от большинства m, с которым оно вынесено, но никак не от общего числа n присяжных. И по существу противное решение таким же количеством голосов при равенстве шансов ошибок всех присяжных не может ни повысить, ни понизить основание полагать, что [общее] решение верно или ошибочно. Но этот вывод существенно предполагает, что шанс u безошибочности присяжных известен до их решения. Это не будет иметь места, как указано ниже, когда он выводится после решения по числу голосов за и против.

При заданном значении u решение, принятое при нечётном числе присяжных большинством, к примеру, всего в один голос, не заслуживает ни более, ни менее доверия, чем решение одного-единственного присяжного. Однако, вероятность подобного решения, будь то осуждения или оправдания, убывает по мере возрастания общего числа присяжных. Эта вероятность равна сумме правых частей формул (117.1) и (117.2) при $n = 2i + 1$. Обозначив эту сумму через w_i и учитывая значение N_i, мы имеем

$$w_i = \frac{(2i+1)![u(1-u)]^i}{(i!)^2(i+1)},\ w_{i+1} = \frac{2i+3}{2i+4}4u(1-u)w_i.$$

Поскольку $4u(1 - u)$ не может превзойти единицу, неизменно будет $w_{i+1} < w_i$. Сравнивая это значение w_i с H_i из § 118, мы видим, что первое превосходит второе в отношении $1:2u(1 - u)$, каково бы ни было i.

120. Если известно только, что подсудимый осуждён большинством по меньшей мере m голосов, так что оно может быть равным m, $m + 2$, $m + 4$, …, $m + 2i$, или единогласным, оказывается, что вероятность его вины будет выше, чем p_i; я обозначу её через P_i. Предполагая, что осуждаемый виновен, вероятность осуждения подсудимого, или наблюдённого события, в соответствии со сказанным выше равна kU_i. Она окажется равной $(1 - k)V_i$, если предполагать, что он невиновен. Поэтому, добавляя вероятность оправдания невиновности [тем же] большинством по меньшей мере в m голосов, мы имеем

$$P_i = \frac{kU_i}{kU_i + (1-k)V_i},\ Q_i = \frac{(1-k)U_i}{(1-k)U_i + kV_i}. \qquad (120.1,\ 2)$$

Вероятности верности решения указанным большинством будут также выражены величинами P_i и Q_i. Они[3], в отличие от p_i и q_i, не будут независимы от общего числа n присяжных и зависят только от m или $n - 2i$. Для количественного сравнения одних с другими я приму $k = 1/2$, что приравняет величины P_i и Q_i равно как и p_i и q_i, и предположу, что до решения дела невиновность подсудимого столь же вероятна, как его вина. Наконец, я приму, что $u = 3/4$, так что можно ставить 3 против одного за то, что ни один присяжный не ошибся.

При обычном числе присяжных, т. е. при $n = 12$, и при $i = 5$ я прежде всего нахожу, что $p_i = 9/10$ и $1 - p_i = 1/10$ и далее, что $U_i = 7254 \cdot 3^7/4^{12}$ и $V_i = 239\ 122/4^{12}$. При этом оказывается, что почти точно $P_i = 403/409$, $1 - P_i = 6/409$.

Это доказывает, что в приведенном примере вероятность $1 - P_i$ ошибочного осуждения большинством по меньшей мере в два голоса составляет едва седьмую вероятности $(1 - p_i)$ ошибки, которой можно опасаться при решении большинством в два голоса или семью голосами против пяти.

Формулы (117.1, 117.2; 118.1; 119.1, 119.2; 120.1,120.2) без труда применимы к случаю, при котором подсудимый, предстающий перед тем же числом присяжных, был уже осуждён или оправдан другим судом. В этом случае значением k, которое включено в указанные формулы, является вероятность вины подсудимого, вытекающая из первого решения и определяемая по одной из перечисленных формул.

121. Если $(n - i)$ и i очень большие числа, при вычислении значений U_i и V_i необходимо прибегать к приближённым методам. Для этой цели я принимаю $1 - u = v$, и величина U_i оказывается суммой первых $(i + 1)$ членов разложения $(u + v)^n$, расположенного по возрастающим степеням v. В соответствии с формулами (77.1a, b), при замене p, q, n, μ на u, v, i, n мы получим

$$U_i = \frac{1}{\sqrt{\pi}}\int_\theta^\infty \exp(-t^2)dt + \frac{(n+i)\sqrt{2}}{3\sqrt{\pi n i(n-i)}}\exp(-\theta^2), \qquad (121.1a)$$

$$U_i = 1 - \frac{1}{\sqrt{\pi}}\int_\theta^\infty \exp(-t^2)dt + \frac{(n+i)\sqrt{2}}{3\sqrt{\pi n i(n-i)}}\exp(-\theta^2), \qquad (121.1b)$$

причём квадрат положительной величины θ равен

$$\theta^2 = i\ln\frac{i}{v(n+1)} + (n+1-i)\ln\frac{n+1-i}{u(n+1)}.$$

Первое и второе из приведенных выражений для U_i соответствуют случаям, при которых v/u превышает отношение $i/(n + 1 - i)$ или меньше него.

Если подсудимый был осуждён при любом большинстве от наименьших, в один или два голоса, или единогласно, то число $(n + 1)$ и указанное отношение будут почти равны $2i$ и 1, а при $v > u$ и $v < u$ следует применять первую или вторую из формул (121.1). А если u и v значительно отличаются от 1/2, или $4uv$ от единицы, то почти точно $\theta = i\ln(1/4uv)$.

Итак, поскольку i – очень большое число, значение θ^2 окажется достаточно большим, чтобы интегралы и экспоненциальные функции в формулах (121.1) стали неощутимыми. Величина U_i будет равна единице или нулю при $u > v$ и $u < v$. А так как в нашем случае сумма $U_i + V_i$ равна почти точно или в точности единице при нечётном и чётном i, то V_i будет равна нулю или единице при $u > v$ и $u < v$.

Отсюда следует, что если предварительная вероятность k вины подсудимого не будет очень небольшой дробью и если шанс v ошибки каждого присяжного заметно меньше шанса u противного, то вероятность P_i его вины после осуждения очень большим числом n присяжных будет очень близка достоверности. Этот результат вызван тем, что при указанном большом числе n очень маловероятно, чтобы решение выносилось слабым большинством. Напротив, если u ощутимо меньше v, а k не является дробью, очень близкой единице, вероятность P_i верности решения станет весьма малой дробью, а невиновность подсудимого окажется крайне вероятной. Вероятности осуждения и оправдания, c_i и d_i, указанные формулами (118.1), станут очень мало отличаться от k и $(1 - k)$ когда $u > v$ и, напротив, от $(1 - k)$ и k, если $v > u$.

В случае $u = v = 1/2$ и $n = 2i + 1$ или $2i + 2$ отношение $i/(n + 1 - i)$ станет немного меньше, чем v/u или единицы, так что нужно будет применять формулу (121.1а), И поскольку θ оказывается очень малой дробью, то, если не учитывать ни квадрата θ, ни членов с i в знаменателе, и заметить, что в формулах (121.1)

$$\int_\theta^\infty \exp(-x^2)dx = \frac{\sqrt{\pi}}{2} - \theta,$$

то равенство

$$U_i = \frac{1}{2} + \frac{1}{\sqrt{\pi i}} - \frac{\theta}{\sqrt{\pi}}$$

окажется почти точным.

При нечётном $n = 2i + 1$, подставляя 1/2 вместо u и v, мы получим

$$\theta^2 = -i \ln[1 + \frac{1}{i}] - (i + 2) \ln[1 - \frac{1}{i + 2}].$$

Раскладывая логарифмы в ряды, можно вывести с принятой нами точностью, что $\theta = 1/\sqrt{i}$, $U_i = 1/2$. Сумма $V_i + U_i = 1$, так что V_i также равна 1/2 и, как это и должно быть, окажется, что $P_i = k$, каково бы ни было число присяжных, если их шансы ошибаться и не ошибаться равны. При $n = 2i + 2$, $u = v = 1/2$ будет

$$\theta^2 = -i \ln[1 + \frac{3}{2i}] - (i + 3) \ln[1 - \frac{3}{2i + 6}], \ \theta = \frac{3}{2\sqrt{i}}, \ U_i = \frac{1}{2} - \frac{1}{2\sqrt{\pi i}}.$$

Но в соответствии со значением H_i из § 118 в этом случае будет

$$U_i + V_i = 1 - \frac{1}{\sqrt{\pi i}}.$$

Однако,

$$V_i = \frac{1}{2} - \frac{1}{2\sqrt{\pi i}}$$

и, поскольку $U_i = V_i$, то, как и в предыдущем случае $P_i = k$. Вероятность c_i осуждения, которую мы рассматриваем, будет независима от k и равна U_i или окажется чуть ниже 1/2.

122. Неизменно предполагая, что суд состоит из некоторого числа n присяжных, представим теперь, что у каждого из них шанс не ошибаться может принимать v различных не равновероятных значений. Пусть $x_1, x_2, \ldots, x_v$ будут этими шансами у первого присяжного, $x'_1, x'_2, \ldots, x'_v$ у второго, $x''_1, x''_2, \ldots, x''_v$ у третьего и т. д. Обозначим вообще, далее, через X_i, X'_i,

X''_i, … вероятности того, что шансы x_i, x'_i, x''_i … будут иметь место. Они же будут вероятностями соответствующих шансов $1 - x_i$, $1 - x'_i$, $1 - x''_i$, … Один из шансов $x_1, x_2, \ldots, x_\nu$ наверняка произойдёт, и то же самое будет иметь место по отношению к одному из шансов $x'_1, x'_2, \ldots, x'_\nu$ и $x''_1, x''_2, \ldots, x''_\nu$ и т. д., так что

$$\begin{aligned} &X_1 + X_2 + \ldots + X_\nu = 1, \\ &X'_1 + X'_2 + \ldots + X'_\nu = 1, \\ &X''_1 + X''_2 + \ldots + X''_\nu = 1, \ldots \end{aligned} \tag{121.1}$$

Если принять, что

$$\begin{aligned} &X_1 x_1 + X_2 x_2 + \ldots + X_\nu x_\nu = u, \\ &X'_1 x'_1 + X'_2 x'_2 + \ldots + X'_\nu x'_\nu = u', \\ &X''_1 x''_1 + X''_2 x''_2 + \ldots + X''_\nu x''_\nu = u'', \end{aligned} \tag{121.2}$$

то в то же время

$$\begin{aligned} &X_1(1 - x_1) + X_2(1 - x_2) + \ldots + X_\nu (1 - x_\nu) = 1 - u, \\ &X'_1(1 - x'_1) + X'_2(1 - x'_2) + \ldots + X'_\nu (1 - x'_\nu) = 1 - u', \\ &X''_1(1 - x''_1) + X''_2(1 - x''_2) + \ldots + X''_\nu (1 - x''_\nu) = 1 - u'' \end{aligned}$$

и u, u', u'' и $(1 - u)$, $(1 - u')$, $(1 - u'')$ будут средними значениями шансов не ошибаться и ошибаться у соответствующих присяжных.

В таком случае вероятность П того, что не ошибается никто из имеющих шансы x_i, x'_i, x''_i … не ошибаться, окажется равной произведению этих шансов на их соответствующие вероятности X_i, X'_i, X''_i, …:

$$\Pi = X_i X'_i X''_i \ldots x_i x'_i x''_i \ldots$$

Пусть через P обозначена вероятность того, что не ошибается никто из n присяжных, имеющих указанные шансы не ошибаться, каковы бы ни были вероятности их возможных шансов. По правилу § 10 P будет суммой $n\nu$ значений П, получаемой при последовательной подстановке чисел 1, 2, …, ν вместо каждого из n чисел i, i', i'', … Нетрудно заметить, что при любых n и ν эта сумма равна произведению n средних u, u', u'', …, т. е. $P = uu'u'' \ldots$

По отношению к шансам x, x'_i, x''_i … не ошибаться вероятность того, что ошибётся один-единственный присяжный, вычисляется при замене x_i на $(1 - x_i)$ в произведении П, если им был первый присяжный, x'_i на $(1 - x'_i)$, если им был второй и т. д. Назовём Π' полной вероятностью того, что ошибся один-единственный присяжный:

$$\Pi' = X_i X'_{i'} X''_{i''} \ldots [(1 - x_i)x'_{i'} x''_{i''} \ \ldots + x_i(1 - x'_{i'})x''_{i''} \ \ldots + x_i x'_{i'} (1 - x''_{i''}) \ \ldots + \ldots].$$

С учётом всех возможных шансов не ошибаться для каждого присяжного обозначим теперь через P' вероятность ошибаться только одному присяжному. Она окажется сумой $n\nu$ значений Π', полученной при последовательной подстановке всех чисел 1, 2, …, ν в каждый из индексов i, i', i'', … Нетрудно заметить, что эта сумма зависит лишь от средних u, u', u'', … и равна

$$P' = (1 - u)u'u'' \ldots + u(1 - u')u'' \ldots + uu'(1 - u'') \ldots + \ldots$$

Продолжая подобные преобразования, можно придти к следующему общему предложению: Вероятность того, что из n присяжных $(n - i)$ не ошибаются, а i ошибаются, не изменится, если шанс не ошибаться может принимать лишь одно-единственное значение для каждого из них, т. е. u для первого присяжного, u' – для второго, … Поэтому, если средние шансы u, u', u'', … не равны друг другу, различные вероятности осуждения заданным большинством и вины осуждённого определяются по правилам § 116, обобщённым на некоторое число n присяжных. Если же они равны друг другу, то их вероятности выражают формулы (117.1, 117.2; 118.1; 119.1, 119.2; 120.1, 120.2) при подстановке в них среднего шанса, общего для всех присяжных, вместо u.

Можно отчётливо представить себе, что каждому присяжному присуще большое число не равновероятных шансов безошибочности суждения. Для этого следует принять, что список, из которого каждый присяжный должен быть отобран, разделён на ν классов, так что все лица, попавшие в какой-либо класс, обладают одним и тем же шансом не ошибаться. Пусть для списка, из которого должен быть отобран первый присяжный, x_i будет этим шансом одного из классов, а X_i – отношением числа лиц этого класса к общему числу лиц в списке.

Тогда шанс не ошибаться у первого присяжного, если он принадлежит указанному классу, будет равен x_i, а X_i – вероятности его принадлежности, т. е. его шанса x_i. Если второй присяжный должен быть отобран из другого списка, а x'_i – шанс не ошибаться для лиц одного из классов и X'_i – отношение числа лиц этого класса к общему числу лиц в списке, то x'_i окажется шансом этого второго присяжного не ошибаться, если он принадлежит указанному классу, и X'_i – вероятность его принадлежности, т. е. его шанса и т. д. для всех остальных присяжных.

Если присяжные для сессии каких-либо ассизов[4] отбираются случайно из их списка всех лиц, которые могли стать присяжными, то до жеребьёвки средние шансы u, u', u'', … равны друг другу. Но их совместное значение может оказаться другим для других ассизов. […] В любом случае нельзя смешивать эти средние шансы, которые существовали до выбора присяжных, с шансами не ошибаться, которые относятся к случайно отобранным присяжным. К этому существенному отличию мы тотчас же вернёмся.

123. Если число ν возможных шансов x_1, x_2, … бесконечно, вероятность каждого окажется бесконечно малой. Пусть Xdx будет вероятностью того, что шанс не ошибаться для присяжного, случайно отобранного из данного списка, равен x. Пусть также u окажется средним из всех возможных шансов с учётом их соответствующих вероятностей. Сумма[5], которая должна равняться единице, и та, которая образует u, становятся в соответствии с предыдущим определёнными интегралами с пределами, равными 0 и 1:

$$\int_0^1 Xdx = 1, \int_0^1 xXdx = u. \qquad (123.1, 2)$$

Положительная величина X может быть непрерывной или разрывной функцией x, совершенно произвольной, но подчиняющейся уравнению (123.1). Каждому данному X будет следовать полностью определённое численное значение u, соответствующее, однако, бесконечному множеству различных выражений X или различных законов вероятностей.

Если все значения x от 0 до 1 равновозможны, X окажется независимым от x, притом равным единице, чтобы удовлетворить уравнению (123.1). В этом случае ввиду уравнения (123.2) $u = 1/2$. Если в указанном интервале X возрастает так, что шанс присяжного не ошибаться становится сам по себе тем вероятнее, чем он ближе к достоверности, и если к тому же X возрастает равномерно[6], то

$$X = \alpha x + \beta$$

при положительных константах α и β. Тогда

$$\int_0^1 Xdx = \alpha/2 + \beta = 1,\ \beta = 1 - \alpha/2,\ X = 1 - \alpha/2 + \alpha x,$$

что исключает неравенство $\alpha > 2$. В результате

$$u = 1/2 + \alpha/12$$

и средний шанс не может ни превосходить 2/3, ни быть меньшим, чем 1/2; эти крайние случаи соответствуют значениям $\alpha = 2$ и 0.

Предположим ещё, что X изменяется по геометрической прогрессии при равных приращениях x и удовлетворяет притом уравнению (123.1) и примем, что

$$X = \frac{\alpha}{e^{\alpha} - 1} e^{\alpha x}$$

при произвольном значении α. Как обычно, e означает основание неперовых логарифмов. Мы получим

$$u = \frac{1}{1 - e^{-\alpha}} - \frac{1}{\alpha}.$$

Следовательно, если α возрастает от $-\infty$ до ∞, средний шанс u сможет принимать все возможные значения от 0 до 1: при $\alpha = -\infty$, 0, ∞ соответственно окажется, что $u = 0$, 1/2 и 1.

Пусть различные шансы не ошибаться не изменяются от 0 до 1, а заключены в более тесные пределы; к примеру, шанс x не опускается ниже 1/2; и, кроме того, если он выше, все его значения равновозможны. Тогда X должен быть разрывной функцией, которая определяется следующим образом. Я обозначу через ε положительную и конечную, но совсем неощутимую величину, и пусть fx будет функцией, очень быстро изменяющейся в пределах $x = (1/2 - \varepsilon)$ и 1/2, исчезающей в пределах $x = 0$ и $(1/2 - \varepsilon)$ и принимающей данное постоянное значение g в пределах $x = 1/2$ и 1.

При этих условиях я принимаю, что $X = fx$; по природе этой функции fx окажется, что

$$\int_0^1 Xdx = g/2 + \int_{1/2-\varepsilon}^{1/2} fxdx.$$

Но ввиду условия (123.1)

$$\int_{1/2-\varepsilon}^{1/2} fxdx = 1 - g/2,$$

так что $g \leq 2$, потому что fx может быть только положительной.

В интеграле

$$\int_{1/2-\varepsilon}^{1/2} xfxdx$$

мы можем считать сомножитель x константой, равной 1/2, и интеграл поэтому будет равен $1/2 - g/4$. И, замечая, что

$$\int_0^1 xfxdx = \int_0^{1/2-\varepsilon} xfxdx + \int_{1/2-\varepsilon}^{1/2} xfxdx + \int_{1/2}^1 xfxdx,$$

мы заключаем, что указанная сумма равна $1/2 - g/4 + g/2 - g/8$, т. е. $u = 1/2 + g/8$.

Следовательно, в этом случае средний шанс не может ни превзойти $u = 3/4$, что соответствует значению $g = 2$, ни быть меньшим, чем 1/2, что соответствует значению $g = 0$.

Таким образом, о виде функции X можно принять бесконечное множество различных гипотез. Если одна из них достоверна, соответствующее значение u также окажется несомненным. Если же, напротив, все они возможны, их соответствующие вероятности окажутся бесконечно малыми, и то же произойдёт с различными значениями среднего шанса, который появится ввиду этих гипотез. Последний случай имеет место, когда различные значения, которые может принимать шанс какого-либо присяжного не ошибаться, остаются неизвестными, и мы даже не знаем закона их вероятностей. Об этом законе можно сформулировать все возможные предположения, так что средний шанс примет не равновероятные значения. Обозначим через $\varphi u du$ бесконечно малую вероятность того, что этот шанс в точности равен u. Функция φ может быть непрерывной или разрывной, интеграл от неё в пределах 0 и 1 должен быть равен единице, и к ней относятся все замечания, приведенные по поводу X.

124. Предыдущие формулы полностью отвечают на все вопросы, относящиеся к теме этой главы, если только известна предварительная вероятность k вины подсудимого, а для каждого присяжного и каждого вида дел известна вероятность не ошибаться. Если же этот шанс принимает многие возможные значения, то требуется, чтобы все эти значения, равно как и их вероятности, были заданы. Опять же, если этих значений бесконечно много, а вероятность каждого бесконечно мала, необходимо знать функцию, которая выражает закон их вероятностей.

Но ни один из этих необходимых элементов заранее не известен. До появления подсудимого перед присяжными привлечение к суду и соответствующее судопроизводство несомненно приводят к вероятности его вины, превышающей вероятность невиновности. Поэтому имеются основания полагать, что $k > 1/2$, но насколько? Мы никак не можем знать этого заранее. Всё зависит от умения и суровости магистратов, которым было поручено предварительное следствие и может изменяться с изменением категории дел. И мы также не можем узнать ни до, ни после отбора присяжных по жребию их шансы не ошибаться. Этот шанс для каждого из них зависит от его просвещённости, целесообразности, которую он придаёт борьбе с тем или иным видом преступности, жалости, которую в нём возбуждает возраст или пол подсудимого и т. д. Ни одно из этих обстоятельств нам не известно, неизвестно и их количественное влияние на голосование. Чтобы можно было применять предыдущие формулы, необходимо исключить из них неизвестные элементы, и этим мы сейчас и займёмся.

125. Рассмотрим случай, при котором шанс не ошибаться один и тот же у всех присяжных. Предположим, что до судебного решения он неизвестен, может принимать все возможные значения от 0 до 1, причём бесконечно малая вероятность его значения, равного u, обозначена через $\varphi u du$. Если это значение достоверно, т. е. если шанс не ошибаться действительно равен u для каждого присяжного, формула (117.1) выразит вероятность того, что подсудимый, будь он виновен или нет, будет осуждён

($n-i$) голосами против i других. В этой формуле n – общее число присяжных, а k – предварительная вероятность вины подсудимого. Следовательно, вероятность подобного разделения голосов будет действительно определяться правой частью формулы (117.1), умноженной на $\varphi u du$. И если указанное разделение произошло, вероятность, что единый для всех шанс не ошибаться равен u, окажется упомянутым произведением, делённым на сумму значений этого же произведения при всех значениях u от 0 до 1 (§ 43). Таким образом, обозначив через $w_i du$ эту бесконечно малую вероятность, мы получим

$$w_i = \frac{[ku^{n-i}(1-u)^i + (1-k)u^i(1-u)^{n-i}]\varphi u}{\int_0^1 [ku^{n-i}(1-u)^i + (1-k)u^i(1-u)^{n-i}]\varphi u du}.$$

Множитель N_i в формуле (117.1), не зависящий от u и общий для числителя и знаменателя w_i, здесь сокращён. Если λ_i – вероятность того, что шанс u не ошибаться заключён в заданных границах l и l', эта величина окажется равной

$$\lambda_i = \frac{k\int_l^{l'} u^{n-i}(1-u)^i \varphi u du + (1-k)\int_l^{l'} u^i(1-u)^{n-i}\varphi u du}{k\int_0^1 u^{n-i}(1-u)^i \varphi u du + (1-k)\int_0^1 u^i(1-u)^{n-i}\varphi u du}. \qquad (125.1)$$

При чётном n и разделении голосов поровну $n = 2i$ и, следовательно,

$$\lambda_i = \frac{\int_l^{l'} u^i(1-u)^i \varphi u du}{\int_0^1 u^i(1-u)^i \varphi u du}.$$

Таким образом, λ_i не зависит от k, от которого она вообще зависит при неравном распределении голосов. Если два каких-либо значения u, равно удалённых от крайних значений 0 и 1 или от среднего $u = 1/2$, равновероятны, так что $\varphi(1-u) = \varphi u$, то

$$\int_0^1 u^{n-i}(1-u)^i \varphi u du = \int_0^1 u^i (1-u)^{n-i} \varphi u du.$$

Если кроме того $l < 1/2$ и $l' = 1 - l$, то также

$$\int_l^{l'} u^{n-i}(1-u)^i \varphi u du = \int_l^{l'} u^i (1-u)^{n-i} \varphi u du,$$

а формула (125.1) принимает вид

$$\lambda_i = \frac{\int_l^{1-l} u^{n-i}(1-u)^i \varphi u du}{\int_0^1 u^{n-i}(1-u)^i \varphi u du}$$

и снова оказывается независимой от k, каково ни было бы разделение голосов $(n - i)$ и i.

Если в формуле (125.1) принять $l = 1/2$ и $l' = 1$ и обозначить результат через λ'_i, то

$$\lambda'_i = \frac{k\int_{1/2}^1 u^{n-i}(1-u)^i \varphi u du + (1-k)\int_{1/2}^1 u^i (1-u)^{n-i} \varphi u du}{k\int_0^1 u^{n-i}(1-u)^i \varphi u du + (1-k)\int_0^1 u^i (1-u)^{n-i} \varphi u du}$$

окажется вероятностью того, что шанс u заключен в границах 1/2 и 1, т. е. превосходит половину. Если теперь принять $l = 0$ и $l' = 1/2$, и обозначить соответствующий вид формулы (125.1) через λ''_i, то вероятность того, что u меньше 1/2 равна [повторена формула (125.1), пределы интегралов в числителе которой стали равными 1/2 и 1]. И, поскольку вероятность того, что в точности $u = 1/2$, бесконечно мала, $\lambda'_i + \lambda''_i = 1$. Это можно тотчас же проверить, заметив, что знаменатели обоих величин совпадают, а сумма интегралов в числителе, умноженных на k, равна интегралу в знаменателе, умноженному на k и то же имеет место для интегралов, умноженных на $(1 - k)$.

126. Если шанс не ошибаться у каждого присяжного равен u, вероятность того, что подсудимый признан виновным $(n - i)$

голосами против i, выражена величиной $w_i du$, а вероятность того, что подсудимый притом виновен, если указанные шансы достоверно равны u, выражена величиной p_i из § 119. По правилам §§ 5 и 10 полная вероятность вины подсудимого равна интегралу от $p_i w_i du$ в пределах u = 0 и 1, т. е. следующей величине, в которой учтены выражения w_i и p_i:

$$\chi_i = \frac{k\int_0^1 u^{n-i}(1-u)^i \varphi u du}{k\int_0^1 u^{n-i}(1-u)^i \varphi u du + (1-k)\int_0^1 u^i (1-u)^{n-i} \varphi u du}. \quad (126.1)$$

Эта вероятность равна 0 и 1 одновременно с k. Представив её в виде

$$\chi_i = k + \frac{k(1-k)\left(\int_0^1 u^{n-i}(1-u)^i \varphi u du - \int_0^1 u^i (1-u)^{n-i} \varphi u du\right)}{k\int_0^1 u^{n-i}(1-u)^i \varphi u du + (1-k)\int_0^1 u^i (1-u)^{n-i} \varphi u du},$$

можно заметить, что при всех других значениях k вероятность вины подсудимого после решения дела станет выше или ниже предшествовавшей в соответствии с тем, будут ли первые интегралы в числителе этой формулы больше или меньше вторых. Если же они равны друг другу, что происходит всегда при $n = 2i$ и $\varphi(1 - u) = \varphi u$, то $\chi_i = k$. Действительно, вероятность вины подсудимого не может никак измениться решением при разделении голосов поровну или такого, при котором равновероятны значения u и $(1 - u)$ или $1/2 \pm (1 - 2u)/2$ шанса не ошибаться.

Во всех остальных случаях χ_i зависит, не только от большинства m или $(n - 2i)$ при голосовании и от величины k, как p_i, но и от общего числа присяжных n и закона вероятностей шансов не ошибаться, выраженного функцией φu.

Пусть, к примеру, осуждение объявлено большинством в один-единственный голос при 201 присяжных, или же в другом деле одним-единственным присяжным, притом наверняка шансы присяжных не ошибаться были одними и теми же у всех упомянутых лиц. Вероятности верности решений суда в обоих случаях окажутся в точности равными, однако, если в первом из них этот шанс значительно отклоняется от 1/2, наблюдённое событие окажется исключительным, очень маловероятным и крайне редко происходящим фактом. Если же этот шанс не ошибаться равен половине, то, в соответствии с выражением для w_i из § 119, вероятность первого события будет немного выше 1/9. Но если до решения он неизвестен, и мы устанавливаем его из самого решения, вина подсудимого намного менее вероятна в первом случае, чем во втором. Это объясняется не тем, что сам по себе результат голосования 101:100 менее верен, чем решение единственного присяжного, а тем, что разделение 201 голосов почти поровну означает, что шанс присяжных не ошибаться весьма вероятно мало отличался от половины, несомненно ввиду трудности рассматриваемого дела.

127. Чтобы сформулировать точную идею о смысле формул (125.1) и (126.1), следует предположить, что перед решением дела некто имел основания полагать подсудимого виновным с вероятностью k, что он не знаком с рассматриваемым делом, не знает никого из n присяжных, разве только, что они были случайно отобраны из общего списка. Для него вероятность присяжного не ошибаться при голосовании одна и та же для всех них (§ 122), но неизвестна. Перед решением можно предположить, что это неизвестное u приняло любое возможное значение от 0 до 1. По каким-то соображениям, которые мы оставляем совершенно в стороне, бесконечно малая вероятность того, что указанное лицо придаёт переменной u, равна $\varphi u du$ при данной функции φu. Интеграл от неё в пределах 0 и 1 должен равняться единице, потому что значение u несомненно содержится в тех же границах.

После вынесения решения известно, что подсудимый был оправдан i голосами и осуждён $(n - i)$ из них. Это – дополнительное сведение, и для упомянутого лица вероятность λ_i что шанс каждого присяжного не ошибаться, равный u, будет теперь заключаться в границах l и l'. Основание полагать подсудимого виновным также усиливается или ослабляется, потому что вероятность k, существовавшая до решения, стала равной χ_i. Для другого лица, которому известны другие данные, это основание окажется иным, но его нельзя смешивать с самим шансом вины. Последний зависит от k и шанса не ошибаться, присущего каждому присяжному, участвующему в голосовании, в зависимости от его способностей и сути рассматриваемого дела. Будь численные значения u, u', u'', … этого шанса известны для каждого присяжного, равно как и k, действительный шанс вины подсудимого после решения дела можно было бы вычислить по правилам § 116, обобщённым на случай n присяжных. Но ввиду невозможности априорного установления этих значений указанные правила невозможно применить.

Если известно лишь, что подсудимый осуждён большинством по меньшей мере в m или $(n - 2i)$ голосов, которое могло быть равным $m + 2$, $m + 4$, … или же осуждён единогласно, и если в этом случае обозначить через Y_i вероятность того, что шанс не ошибаться, единый для всех присяжных, заключён в границах l' и l, и через Z_i вероятность того, что осуждённый виновен, выражения для этих двух величин можно будет получить при помощи того же самого рассуждения, которое было использовано для вывода λ_i и χ_i, но применяя c_i и P_i (§§ 118 и 120) вместо γ_i и p_i, как при установлении формул (125.1) и (126.1). Так мы получим

$$Y_i = \frac{k\int_0^1 U_i \varphi u du + (1-k)\int_l^{l'} V_i \varphi u du}{k\int_0^1 U_i \varphi u du + (1-k)\int_0^1 V_i \varphi u du}, \tag{127.1a}$$

$$Z_i = \frac{k\int_0^1 U_i \varphi u du}{k\int_0^1 U_i \varphi u du + (1-k)\int_0^1 V_i \varphi u du}. \tag{127.1b}$$

Эти формулы, как и формулы (125.1) и (126.1), можно обобщить на случай, при котором известно, что n' присяжных из n были отобраны по жребию из первого списка, n'' – из другого списка и т. д., если также предположить, что в первом списке средний шанс u' не ошибаться имел вероятность $\varphi' u' du'$, во втором списке соответственно u'' и $\varphi'' u'' du''$ и т. д. Впрочем, это обобщение и нетрудно, и бесполезно, и мы не станем выписывать сложные формулы, которые имели бы место[7].

128. Если i и $(n - i)$ очень большие числа, следует применить метод § 67 для вычисления приближённых значений интегралов, содержащихся в формулах (125.1, 126.1, 127.1). Я начну с формул (125.1) и (126.1).

В интервале $u = 0$ и 1 произведение $u^{n-1}(1 - u)^i$ имеет лишь один максимум. Я обозначу его значение через β, а через α – соответствующее значение u. Тогда

$$\alpha = \frac{n-i}{n},\ \beta = \frac{i^i (n-i)^{n-i}}{n^n},\ u^{n-i}(1-u)^i = \beta \exp(-x^2).$$

Далее, логарифмируя, я получу

$$x^2 = \lg\beta - (n-i)\lg u - i\lg(1-u).$$

Переменная x монотонно возрастает от $-\infty$ до ∞ и значения $x = -\infty, 0, \infty$ соответствуют $u = 0, \alpha, 1$. Пределы интеграла по x равны $\pm\infty$ при $u = 0$ и 1. Вообще, обозначая пределы по x, соответствующие пределам l и l' по u, через λ и λ', с учётом предыдущих значений β и x^2 мы получим

$$\lambda = \pm\sqrt{(n-i)\ln[(n-i)/ln] + i\ln[i/n(1-l)]}\,,$$
$$\lambda' = \pm\sqrt{(n-i)\ln[(n-i)/l'n] + i\ln[i/n(1-l')]}.$$

Если l и l' превышают α, то λ и λ' станут положительными, и перед радикалами следует принять верхние знаки, в противном же случае – нижние знаки. А если $l < \alpha$ и $l' > \alpha$, то надо будет принять верхний знак перед вторым радикалом и нижний перед первым, чтобы λ оказалось отрицательным, а λ' положительным.

Чтобы выразить u рядом, расположенным по степеням x, примем, что коэффициенты $\gamma, \gamma', \gamma'', \ldots$ постоянны и положим, что

$$u = \alpha + \gamma x + \gamma' x^2 + \gamma'' x^3 + \ldots$$

Учитывая значения α, β и x^2, мы получим

$$x^2 = \frac{n^3}{2i(n-i)}(\gamma x + \gamma' x^2 + \gamma'' x^3 + \ldots)^2 +$$
$$\frac{n^4(n-2i)}{3i^2(n-i)^2}(\gamma x + \gamma' x^2 + \gamma'' x^3 + \ldots)^3 + \ldots$$

Приравнивая коэффициенты при одних и тех же степенях x в обеих частях этого уравнения [!], можно вывести значения γ, γ', γ'', …, так что

$$u = \alpha + x\sqrt{\frac{2i(n-i)}{n^3}} - \frac{2x^2(n-2i)}{3n^2} + \ldots,$$

$$du = \sqrt{\frac{2i(n-i)}{n^3}}dx - \frac{4x(n-2i)}{3n^2} + \ldots$$

Если функция φu не убывает очень быстро при изменении u в какую-либо сторону от частного значения $u = \alpha$, можно будет после подстановки u, разложенного в ряд, разложить и φu по степеням $(u - \alpha)$, а затем по степеням x. Так мы получим

$$\varphi u = \varphi\alpha + [x\sqrt{\frac{2i(n-i)}{n^3}} - \ldots]\frac{d\varphi\alpha}{d\alpha} + \frac{1}{2}[x\sqrt{\frac{2i(n-i)}{n^3}} - \ldots]\frac{d^2\varphi\alpha}{d\alpha^2} + \ldots$$

При помощи этих различных выражений разложение в ряд интегралов от $u^{n-1}(1 - u)^i\varphi u du$ в пределах 0, 1 из формулы (126.1) будет содержать интегралы в пределах $x = -\infty$ и ∞ от $\exp(-x^2)dx$, умноженного на чётные или нечётные степени x. Значение интегралов, содержащих чётные степени x, известны, остальные же исчезнут. Числа i и $(n - i)$ имеют тот же порядок, что и n, и рассматриваемый ряд будет расположен по величинам порядка малости $1/\sqrt{n}$, $1/(n\sqrt{n})$, $1/(n^2\sqrt{n})$, … Ограничиваясь его первым членом, и заметив, что

$$\int_{-\infty}^{\infty} \exp(-x^2)dx = \sqrt{\pi},$$

мы установим, что

$$\int_0^1 u^{n-i}(1-u)^i \varphi u du = \frac{i^i(n-i)^{n-i}\sqrt{2\pi i(n-i)}}{n^{n+1}\sqrt{n}}\varphi(\frac{n-i}{n}).$$

Кроме того, поменяв местами i и $(n - i)$, мы получим

$$\int_0^1 u^i(1-u)^{n-i} \varphi u du = \frac{i^i(n-i)^{n-i}\sqrt{2\pi i(n-i)}}{n^{n+1}\sqrt{n}}\varphi(\frac{i}{n}).$$

Если обозначить через δ положительную величину, очень малую по сравнению с $\sqrt{n}$, положить

$$l = \frac{n-i}{n} - \delta\sqrt{\frac{2i(n-i)}{n^3}},\ l' = \frac{n-i}{n} + \delta\sqrt{\frac{2i(n-i)}{n^3}}$$

и, пренебрегая членами порядка малости $1/\sqrt{n}$, разложить в ряд логарифмы в выражениях для λ и λ', окажется, что $\lambda = -\delta$ и $\lambda' = \delta$. После этого мы получим

$$\int_l^{l'} u^{n-i}(1-u)^i \varphi u du = \frac{i^i(n-i)^{n-i}\sqrt{2i(n-i)}}{n^{n+1}\sqrt{n}} \varphi(\frac{n-i}{n}) \int_{-\delta}^{\delta} \exp(-x^2)dx$$

с точностью до членов порядка $1/n$. По мере возрастания δ интеграл в правой части будет стремиться к $\sqrt{\pi}$. Для его очень небольшого отличия от этого значения достаточно, чтобы δ была равна таким числам, как 2 или 3. При пределах l и l', ставших заметно больше или меньше $(n-i)/n$, интеграл в левой части равенства почти исчезнет.

Обозначим через ε положительную величину, очень малую по сравнению с $\sqrt{n}$, и положим, что

$$l = \frac{i}{n} - \varepsilon\sqrt{\frac{2i(n-i)}{n^3}},\ l' = \frac{i}{n} + \varepsilon\sqrt{\frac{2i(n-i)}{n^3}}\ .$$

Тогда

$$\int_l^{l'} u^i(1-u)^{n-i} \varphi u du =$$

$$\frac{i^i(n-i)^{n-i}\sqrt{2i(n-i)}}{n^{n+1}\sqrt{n}} \varphi(\frac{i}{n}) \int_{-\varepsilon}^{\varepsilon} \exp(-x^2)dx\ . \qquad (128.1)$$

При пределах l и l', ставших заметно больше или меньше i/n, интеграл в левой части равенства почти исчезнет.

Если дроби $(n-i)/n$ и i/n заметно отличаются друг от друга, то первые из предыдущих значений l и l' будут также отличаться от значения $u = i/n$, соответствующего максимуму $u^i(1-u)^{n-i}$ и интеграл в левой части уравнения (128.1), равно как и последующие значения l и l', будут также заметно отличаться от значения $u = (n-i)/n$, соответствующего максимуму $u^{n-i}(1-u)^i$, и интеграл в левой части уравнения (128.1) с переставленными показателями степеней также станет почти равным нулю.

129. Подставляя в формулу (126.1) приближённые значения входящих в неё интегралов и сокращая общие множители в числителе и знаменателе, мы получим вероятность вины подсудимого, осуждённого большинством в $m = n - 2i$ голосов при очень большом числе n присяжных

$$\chi_i = \frac{k\varphi[(n-i)/n]}{k\varphi[(n-i)/n] + (1-k)\varphi(i/n)}.$$

Заметно, что она зависит от отношения i/n, или, если угодно, от отношения $(n - i)/n$, но не от разности этих чисел, как вероятность p_i в случае, при котором (§ 119) шанс u не ошибаться был известен заранее. К примеру, если подсудимый осуждён 1000 голосами против 500 при 1500 присяжных или же 100 голосами против 50 при 150 присяжных, вероятности χ_i в этих случаях окажутся теми же самыми, однако вероятности p_i будут весьма отличаться друг от друга. Напротив, при том же втором случае пусть в первом из них 550 присяжных осудили подсудимого и 500 оправдали, тогда вероятности p_i окажутся теми же самыми, а вероятности χ_i могут намного отличаться друг от друга.

Пусть подсудимый осуждён, тогда $(n - i)/n > 1/2$ и $i/n < 1/2$. Предположим, что при $u < 1/2$ функция φu будет почти нулём; иначе говоря, что средний шанс не ошибаться считается неправдоподобным, если он окажется меньшим 1/2 или меньшим шанса ошибаться. И пусть, кроме того, дробь k не слишком близка нулю, тогда можно будет пренебречь вторым членом знаменателя χ_i по сравнению с первым, так что окажется, что вероятность $\chi_i = 1$ или по крайней мер очень близка к достоверности.

Полагаясь на приближённые значения интегралов в формуле (125.1) и предполагая, что дроби $(n - i)/n$ и i/n не слишком мало отличаются друг от друга, мы получим вероятность

$$\lambda_i = \frac{\frac{1}{\sqrt{\pi}} k\varphi(\frac{n-i}{n}) \int_{-\delta}^{\delta} \exp(-x^2)dx}{k\varphi(\frac{n-i}{n}) + (1-k)\varphi(\frac{i}{n})}$$

того, что при решении осудить подсудимого единый шанс u не ошибаться заключен в границы

$$\frac{n-i}{n} \mp \delta\sqrt{\frac{2i(n-i)}{n^3}}.$$

Оставляя ту же гипотезу, при которой один из двух интегралов в числителе формулы (125.1) становится равным нулю, мы имеем

$$\lambda_i = \frac{\frac{1}{\sqrt{\pi}} k\varphi(\frac{i}{n}) \int_{-\varepsilon}^{\varepsilon} \exp(-x^2)dx}{k\varphi(\frac{n-i}{n}) + (1-k)\varphi(\frac{i}{n})}$$

для вероятности, что этот шанс заключён в границы

$$\frac{i}{n} \mp \varepsilon\sqrt{\frac{2i(n-i)}{n^3}}.$$

Величинам δ и ε можно придать достаточно, хотя и не слишком, большие значения, чтобы интегралы с переменной x очень немного отличались от $\sqrt{\pi}$. Тогда сумма обоих значений λ_i также очень мало будет отличаться от единицы. Будет почти достоверно, что средний шанс u окажется либо в первых границах, мало отличающихся от дроби $(n-i)/n$, превышающей половину, либо во вторых границах, мало отличающихся от i/n, которая меньше половины.

Если предположить, что $\varphi(i/n)$ неощутимо или пренебрегаемо по сравнению с $\varphi[(n-i)/n]$, второй случай исключается, и можно считать почти достоверным, что значение u очень немного отличается от $(n-i)/n$. Иначе говоря, что шансы u и $(1-u)$ не ошибаться и ошибаться относятся как количества осуждающих и оправдывающих голосов $(n-i)$ и i. После этого представляется, что вероятность χ_i не сводится ощутимо к единице, а очень мало отличается от p_i при $u = (n-i)/n$. Но следует заметить, что вероятность p_i соответствует случаю, при котором шанс u определённо принимает лишь одно возможное значение. Чтобы включить этот случай в тот, которому соответствует χ_i, следует предположить, что φu равна нулю только в бесконечно малых интервалах по обе стороны от указанного возможного значения u, и что эта функция очень быстро убывает в окрестности этого значения. Но анализ в § 128 существенно предполагал, как это было заметно, что φu нисколько не изменяется ни в какую сторону от значения $(n - i)/n$. Поэтому выведенное там χ_i никак не применимо к случаю, которому соответствует p_i из § 119. Можно также заметить, что последний описан формулой (126.1).

Если вообще обозначить через v единственное возможное значение u и через η положительную бесконечно малую и принять за φu функцию, равную нулю при всех значениях u вне границ $v \pm \eta$, пределы интегралов в формуле (126.1) совпадут с этими граничными точками. Внутри этих пределов множители $u^{n-i}(1-u)^i$ и $u^i(1-u)^{n-i}$ окажутся константами; после вывода их за знак интегралов и сокращения в числителе и знаменателе формулы (126.1) общего множителя, интеграла от $\varphi u du$ в пределах $v - \eta$ и $v + \eta$, эта формула совпадёт с формулой (119.1) при $u = v$.

Если дроби $(n - i)/n$ и i/n не отличаются друг от друга существенно, и если принять $\varepsilon = \delta$, то предшествующие значения λ_i будут относиться к тем же границам шанса u. Но их единое значение будет отличаться от предшествовавших, станут независимыми от k и равными

$$\frac{1}{\sqrt{\pi}}\int_{-\delta}^{\delta}\exp(-x^2)dx,$$

потому что в этом частном случае оба интеграла в числителе формулы (125.1) почти равны друг другу, равно как и интегралы в её знаменателе.

130. Для определения приближённых значений интегралов в формуле (127.1) следует выразить значения U_i и V_i при помощи формул (121.1). Первая из них относится к случаю, при котором $(1 - u)/u$ превосходит $i/(n + 1 - i)$, а вторая – к противному случаю. Первая, стало быть, имеет место в пределах $u = 0$ и α, вторая – в пределах $u = \alpha$ и 1, при замене $(n + 1)$ на n и равенстве $(n - i)/n = \alpha$. В соответствии с уравнением, которое определяет θ в формулах (121.1), окажется, что

$$u^{n-i}(1-u)^i = \frac{i^i(n-i)^{n-i}}{n^n}\exp(-\theta^2).$$

Это уравнение совпадает с указанным [в § 128], которое связывает u и x. Из него следует, что

$$u=\alpha+\theta\sqrt{\frac{2i(n-i)}{n^3}}-\frac{2\theta^2(n-2i)}{3n^2}+\dots$$

Но θ должно всегда быть положительной величиной (§ 121), и при $u = 0$, α и 1 оно равно ∞, 0 и ∞. Переменная u возрастает от 0 до α, а θ убывает от ∞ до 0. Затем u снова возрастает от α до 1, а θ возрастает от 0 до ∞. Поэтому, в соответствии с формулами (121.1),

$$\int_0^\alpha U_i\varphi u du=\frac{1}{\sqrt{\pi}}\int_0^\infty\left(\int_\theta^\infty\exp(-x^2)dx\right)\varphi u\frac{du}{d\theta}d\theta+$$
$$\frac{(n+i)\sqrt{2}}{3\sqrt{\pi ni(n-i)}}\int_{-\infty}^0\exp(-\theta^2)\varphi u\frac{du}{d\theta}d\theta,$$

$$\int_0^1 U_i\varphi u du=\int_\alpha^1\varphi u du-\frac{1}{\sqrt{\pi}}\int_0^\infty\left(\int_\theta^\infty\exp(-x^2)dx\right)\varphi u\frac{du}{d\theta}d\theta+$$
$$\frac{(n+i)\sqrt{2}}{3\sqrt{\pi ni(n-i)}}\int_0^\infty\exp(-\theta^2)\varphi u\frac{du}{d\theta}d\theta.$$

Значения этих интегралов, в том числе двойных, в виде сходящихся рядов по θ можно получить при подстановке под знаками интегралов предыдущего ряда вместо u и его производной вместо $du/d\theta$ и при разложении в ряд также и φu, что предполагает, что эта функция не изменяется очень быстро в одну или другую сторону от частного значения аргумента $u = \alpha$. Если пренебречь членами порядка малости $1/n$, то окажется, что просто

$$u=\alpha+\theta\sqrt{\frac{2i(n-i)}{n^3}},\ \frac{du}{d\theta}=\sqrt{\frac{2i(n-i)}{n^3}},\ \varphi u=\varphi\alpha.$$

Интегралы в этих формулах могут предваряться знаками плюс и минус; плюс следует принять, если переменная θ возрастает и минус в противном случае. Затем знаки интегралов меняются, а их пределы меняются местами. В результате оказывается, что

$$\int_0^\alpha U_i\varphi u du=\varphi\alpha\sqrt{\frac{2i(n-i)}{\pi n^3}}\int_0^\infty\left(\int_\theta^\infty\exp(-x^2)dx\right)d\theta,$$

$$\int_{\alpha}^{1} U_i \varphi u du = \int_{\alpha}^{1} \varphi u du - \varphi\alpha\sqrt{\frac{2i(n-i)}{\pi n^3}}\int_{0}^{\infty}\left(\int_{\theta}^{\infty}\exp(-x^2)dx\right)d\theta.$$

Складывая эти формулы, мы получим

$$\int_{0}^{1} U_i \varphi u du = \int_{\alpha}^{1} \varphi u du.$$

В общем случае, если обозначить через a и a_1 два таких значения u, что $a < \alpha$ и $a_1 > \alpha$ и через b и b_1 – положительные значения θ, соответствующие значениям $u = a$ и a_1, то с нашей степенью приближения[8] окажется, что

$$\int_{a}^{\alpha} U_i \varphi u du = \varphi\alpha\sqrt{\frac{2i(n-i)}{\pi n^3}}\int_{0}^{b}\left(\int_{\theta}^{\infty}\exp(-x^2)dx\right)d\theta,$$

$$\int_{\alpha}^{a_1} U_i \varphi u du = \int_{\alpha}^{a_1} \varphi u du - \varphi\alpha\sqrt{\frac{2i(n-i)}{\pi n^3}}\int_{0}^{b_1}\left(\int_{\theta}^{\infty}\exp(-x^2)dx\right)d\theta.$$

Интегрируя по частям, можно получить […] и поэтому

$$\int_{a}^{\alpha} U_i \varphi u du = \varphi\alpha\sqrt{\frac{2i(n-i)}{\pi n^3}}\left(b_1\int_{b}^{\infty}\exp(-x^2)dx + \frac{1}{2} - \frac{1}{2}\exp(-b^2)\right),$$

$$\int_{\alpha}^{a_1} U_i \varphi u du = \int_{\alpha}^{a_1} \varphi u du -$$

$$\varphi\alpha\sqrt{\frac{2i(n-i)}{\pi n^3}}\left(b_1\int_{b_1}^{\infty}\exp(-x^2)dx + \frac{1}{2} - \frac{1}{2}\exp(-b_1^2)\right).$$

В формулах (121.1) я заменил U_i на V_i и поэтому заменил u на $(1 - u)$ (§ 118). Первая замена имеет место, если $u/(1 - u)$ превосходит $i/(n + 1 - i)$, т. е. в интервале $u = 1 - \alpha$ до 1, принимая при этом n вместо $(n + 1)$ и неизменно полагая, что $\alpha = (n - i)/n$.

Вторая замена относится к случаю, при котором u изменяется от 0 до $1 - \alpha$. Обозначив через θ′ значение θ при замене u на $(1 - u)$ и продолжая пренебрегать членами порядка малости $1/n$, мы вначале получаем […], так что

$$\int_0^1 V_i \varphi u du = \int_0^{1-\alpha} \varphi u du.$$

Далее, если a' и a'_1 являются такими значениями u, что $a' < 1 - \alpha$ и $a'_1 > 1 - \alpha$, и если обозначить через b' и b'_1 положительные значения θ', выведенные из уравнения

$$(1-u)^{n-i} u^i = \frac{i^i (n-i)^{n-i}}{n^3} \exp(-\theta'^2),$$

и соответствующие значениям $u = a'$ и a'_1, то будет также

$$\int_{1-\alpha}^{a'_1} V_i \varphi u du =$$

$$\varphi(1-\alpha)\sqrt{\frac{2i(n-i)}{\pi n^3}} \left(b'_1 \int_{b'_1}^{\infty} \exp(-\theta'^2) d\theta' + \frac{1}{2} - \frac{1}{2} \exp(-b_1'^2) \right),$$

$$\int_{a'}^{1-\alpha} V_i \varphi u du =$$

$$\int_{a'_1}^{1-\alpha} \varphi u du - \varphi(1-\alpha)\sqrt{\frac{2i(n-i)}{\pi n^3}} \left(b' \int_{b'}^{\infty} \exp(-\theta'^2) d\theta' + \frac{1}{2} - \frac{1}{2} \exp(-b'^2) \right).$$

131. Приближённые значения интегралов в формулах (127.1), будучи вычислены, мы получаем выражение

$$Z_i = \frac{k \int_\alpha^1 \varphi u du}{k \int_\alpha^1 \varphi u du + (1-k) \int_0^{1-\alpha} \varphi u du} \qquad (131.1)$$

для вероятности вины осуждённого не менее, чем $(n - i)$ присяжными из их общего очень большого числа n. Отношение $\alpha = (n - i)/n$ превышало 1/2. Если предположить, что при $u < 1/2$ функция φu незаметна или равна нулю, то таким же будет второй интеграл в знаменателе формулы (131.1), и если k – не очень небольшая дробь, то Z_i окажется почти равным единице. А если $\varphi(1-u) = \varphi u$ при всех значениях u, то

$$\int_0^{1-\alpha} \varphi u du = -\int_0^{\alpha} \varphi(1-u)du = \int_{\alpha}^{1} \varphi u du$$

и Z_i становится равным k, что и должно было быть.

Если принять $a = 1 - \alpha$ и $a'_i = \alpha$, соответствующие значения b и b'_1 величин θ и θ′ окажутся равными друг другу. Обозначим их через c и учтём то, чем является α, тогда c будет положительной величиной, определяемой из уравнения

$$(n-i)^i i^{n-1} = i^i (n-i)^{n-i} \exp(-c^2).$$

Мы также получим

$$\int_{1-\alpha}^{\alpha} U_i \varphi u du = \varphi\alpha \sqrt{\frac{2i(n-i)}{\pi n^3}} \left(c\int_c^{\infty} \exp(-x^2)dx + \frac{1}{2} - \exp(-c^2) \right),$$

а потому

$$\int_{1-\alpha}^{\alpha} V_i \varphi u du =$$

$$\varphi(1-\alpha)\sqrt{\frac{2i(n-i)}{\pi n^3}} \left(c\int_c^{\infty} \exp(-x^2)dx + \frac{1}{2} - \exp(-c^2) \right),$$

и вероятность того, что при указанном осуждении единый шанс u присяжных не ошибаться заключён в границах $1 - \alpha$ и α, или между i/n и $(n - i)/n$, равна

$$Y_i = \frac{k\varphi\alpha + (1-k)\varphi(1-\alpha)}{k\int_{\alpha}^{1} \varphi u du + (1-k)\int_0^{1-\alpha} \varphi u du} \times$$

$$\left(c\int_c^{\infty} \exp(-x^2)dx + \frac{1}{2} - \exp(-c^2) \right) \sqrt{\frac{2i(n-i)}{\pi n^3}}.$$

Она очень низка ввиду весьма малого радикала. Поэтому очень вероятно, что шанс u был или больше, чем α или меньше, чем

$1 - \alpha$. Для проверки я приму $a_1 = a'_1 = 1$. Соответствующие значения θ и θ' окажутся равными $b_1 = b'_1 = \infty$, так что

$$\int_{\alpha}^{1} U_i \varphi u du = \int_{\alpha}^{1} \varphi u du - \frac{1}{2}\varphi\alpha\sqrt{\frac{2i(n-i)}{\pi n^3}},$$
$$\int_{1-\alpha}^{1} V_i \varphi u du = \frac{1}{2}\varphi(1-\alpha)\sqrt{\frac{2i(n-i)}{\pi n^3}}.$$

Если из этого последнего интеграла вычесть его предыдущее значение в пределах $(1 - \alpha)$ и α, то окажется, что

$$\int_{\alpha}^{1} V_i \varphi u du = \varphi(1-\alpha)\sqrt{\frac{2i(n-i)}{\pi n^3}}\left(\exp(-c^2) - c\int_{c}^{\infty}\exp(-x^2)dx\right).$$

С учётом интегралов от $U_i\varphi u du$ и $V_i\varphi u du$ в пределах α и 1

$$Y_i = \frac{k\int_{\alpha}^{1}\varphi u du - \left(\frac{1}{2}k\varphi\alpha - (1-k)\varphi(1-\alpha)\Omega\right)}{k\int_{\alpha}^{1}\varphi u du + (1-k)\int_{0}^{1-\alpha}\varphi u du},$$
$$\Omega = [\exp(-c^2) - c\int_{c}^{\infty}\exp(-x^2)dx]\sqrt{\frac{2i(n-i)}{\pi n^3}}$$

окажется вероятностью того, что шанс u заключён в границах $u = \alpha$ и 1, т. е. превышает α. Я также приму, что $a = a' = 0$, тогда $b = b' = \infty$, так что […]

$$Y_i = \frac{(1-k)\int_{0}^{1-\alpha}\varphi u du - \sqrt{\frac{2i(n-i)}{\pi n^3}}\Gamma}{k\int_{\alpha}^{1}\varphi u du + (1-k)\int_{\alpha}^{1-\alpha}\varphi u du},$$
$$\Gamma = \left(\frac{1}{2}(1-k)\varphi(1-\alpha) - k\varphi\alpha[\exp(-c^2) - c\int_{c}^{\infty}\exp(-x^2)dx]\right)$$

будет вероятностью того, что шанс u заключён в границах $u = 0$ и $(1 - \alpha)$, т. е., что он меньше, чем $(1 - \alpha)$. Сумма двух последних значений Y_i почти равна единице, что следует проверить. Если значения φu при $u < 1/2$ равны нулю или почти таковы, последнее значение Y_i окажется очень небольшим, а предыдущее – очень мало отличным от достоверности. В любом случае сумма трёх только что вычисленных значений Y_i равна единице, как это и должно было быть.

132. Если даже число n присяжных очень велико, в соответствии с предыдущим мы обязаны сформулировать гипотезу о виде функции φu или законе вероятностей шансов не ошибаться, чтобы можно было вывести вероятность вины осуждённого $(n - i)$ голосами против i. Более необходимая причина существует в обычных случаях, когда число n не очень значительно.

Та гипотеза, которую по этому поводу принял Лаплас, состояла в предположении, что функция φu равна нулю при всех значениях u, меньших 1/2 и постоянна для всех $u > 1/2$. Она сводится к тому, что каждый шанс не ошибаться, меньший противоположного шанса, считается невозможным, а все остальные шансы не ошибаться считаются равновероятными. Эта гипотеза допустима, потому что удовлетворяет условию равенства единице интеграла от $\varphi u du$ в пределах 0 и 1, притом (§ 123) среднее из возможных значений u, или тот же интеграл от $u\varphi u du$, будет заключено в границах 1/2 и 3/4, и, при $u > 1/2$, зависит от значения φu.

В соответствии с этой гипотезой $\varphi u = 0$ при $u < 1/2$ и постоянно при $u > 1/2$, и потому пределы интеграла в формуле (126.1) суживаются и оказываются равными $u = 0$ и 1/2. Функцию φu можно вывести за знак интеграла, а так как

$$\int_{1/2}^{1} u^i (1-u)^{n-i} du = \int_{0}^{1/2} u^{n-i} (1-u)^i du,$$

то указанная формула принимает вид

$$\chi_i = \frac{k\int_{1/2}^{1} u^{n-i}(1-u)^i du}{k\int_{1/2}^{1} u^{n-i}(1-u)^i du + (1-k)\int_{0}^{1/2} u^{n-i}(1-u)^i du}.$$

Здесь в числителе и знаменателе сокращён постоянный общий множитель φu.

Лаплас вовсе не принял во внимание предварительную вероятность вины подсудимого. Поэтому, для совпадения полученной формулы с его результатом следует предположить, что эта вина не более и не менее вероятна, чем его невиновность, т. е. принять, что $k = 1/2$. Тогда

$$\chi_i = \frac{\int_{1/2}^{1} u^{n-i}(1-u)^i\,du}{\int_{0}^{1} u^{n-i}(1-u)^i\,du},\quad 1-\chi_i = \frac{\int_{0}^{1/2} u^{n-i}(1-u)^i\,du}{\int_{0}^{1} u^{n-i}(1-u)^i\,du}.$$

После интегрирования оказывается, что вероятность невиновности подсудимого, осуждённого большинством присяжных в $(n - 2i)$ из их общего числа n, равна

$$1-\chi_i = \frac{1}{2^{n+1}}[1+\frac{n+1}{1}+\frac{(n+1)n}{2!}+...\ +\frac{(n+1)n...(n-i+2)}{i!}].\quad (132.1)$$

Эту последнюю формулу по существу вывел Лаплас (1816/1886, с. 527). Сумма в скобках состоит из $(i + 1)$ членов и равна единице при $i = 0$. Поэтому вероятность ошибочного единогласного осуждения равна $1/2^{n+1}$. Не принимая значения $k = 1/2$ и полагая $i = 0$, мы получим

$$1-\chi_i = \frac{1-k}{k2^{n+1}-(2k-1)} = \frac{1}{2^{n+1}}[1-\frac{(2k-1)(2^{n+1}-1)}{k2^{n+1}-(2k-1)}],$$

т. е. величине, большей или меньшей, чем $1/2^{n+1}$ при k большем или меньшем половины.

В обычном случае $n = 12$; приняв последовательно $i = 0, 1, 2, \dots, 5$, мы получим из формулы (132.1) дроби

1/8192, 14/8192, 92/8192, 378/8192, 1093/8192, 2380/8192, –

вероятности ошибочного осуждения одиннадцатью присяжными против одного, десятью против двух, …, семью против пяти при 12 присяжных. При наименьшем большинстве вероятность ошибки почти равна 2/7, что означает, что для очень большого числа подсудимых, осужденных таким большинством, весьма вероятно, что 2/7 из них было осуждено ошибочно, а для большинства восьми против четырёх указанная вероятность почти равна 1/8.

Приложим гипотезу Лапласа к формуле (125.1) и обозначим через δ положительную величину, не превышающую 1/2. Полагая $k = 1/2$, $l = 1/2$ и $l' = 1/2 + \delta$, мы получим

$$\lambda_i = \frac{\int_{1/2-\delta}^{1/2+\delta} u^{n-i}(1-u)^i du}{\int_0^1 u^{n-i}(1-u)^i du}$$

для вероятности того, что шанс u не ошибаться, который в соответствии с гипотезой не может опускаться ниже 1/2, заключён в границах 1/2 и 1/2 + δ при осуждении $(n - i)$ голосами против i. Интегрирование не представляет затруднений. При $i = 0$ или единогласии окажется, что

$$\lambda_i = [(1/2) + \delta]^{n+1} - [(1/2) - \delta]^{n+1}.$$

Если принять, к примеру, $n = 12$ и $\delta = 0.448$, то почти точно $\lambda_i = 1/2$, так что можно держать пари на равных, что шанс u заключён в границах 0.5 и 0.948. При $\delta = 1/4$, уже не предполагая, что $i = 0$, мы получим

$$\lambda_i = \frac{1}{4^{n+1}}[3^{n+1} - 1 + \frac{n+1}{1}(3^n - 3) + C_{n+1}^2(3^{n-1} - 3^2) + \ldots$$
$$+ C_{n+1}^i(3^{n-i+1} - 3^i)]$$

для вероятности того, что шанс u заключён в границах 1/2 и 3/4, или что он ближе к 1/2, чем к единице. При $n = 12$ и $i = 5$ он равен 0.915 … Можно ставить чуть более десяти против одного на то, что при наименьшем большинстве этот шанс ниже 3/4.

133. Поскольку формула (132.1) была выведена из другой, в которой шанс u не ошибаться был принят одним и тем же для всех присяжных, он, хоть Лаплас и не упомянул этого, не может быть принят для каждого из них в качестве их действительного шанса. Этой величиной следует считать средний шанс для общего списка, из которого присяжные отбираются по жребию (§ 122). В таком списке несомненно найдутся лица, у которых шанс не ошибаться, по крайней мере в трудных делах, меньше 1/2, т. е.

меньше шанса безошибочности. Гипотеза Лапласа требует, однако, чтобы таких лиц было немного, и, следовательно, чтобы они не помешали среднему шансу безошибочности неизменно быть большим, чем 1/2. Знаменитый геометр также предположил, что все значения этого шанса от 1/2 до 1 равновероятны.

Его единственным обоснованием этих двух предположений было утверждение, что *Мнение судьи сильнее склоняется к истине, чем к ошибке*. Однако, исходя от этого принципа, он только заключил, что значения функции φu, которая выражает закон вероятностей значений среднего шанса, должны быть больше при $u > 1/2$, чем в противном случае. Это условие можно выполнить бесконечным множеством способов, не требуя, ни чтобы $\varphi u = 0$ при $u < 1/2$, ни её постоянства при $u > 1/2$. Рассматриваемая нами гипотеза недостаточно априорно обоснована, а выводимые из неё следствия, как будет показано, делают её совсем неприемлемой.

По существу формула (132.1), которая является одним из её необходимых следствий, не включает ничего, зависящего от способностей лиц из общего списка присяжных. Кто-либо, знающий, к примеру, что два осуждения были вынесены одним и тем же большинством при одном и том же общем числе присяжных, отобранных из двух различных списков, мог бы с тем же основанием полагать, что оба осуждения были ошибочны, хоть и знал, что лица из одного списка намного способнее, чем отобранные из другого списка, и с этим уже невозможно согласиться.

Если подсудимый осуждён большинством $(n - i)$ голосов против i, так что i/n менее половины, то в соответствии с исследуемой нами гипотезой значение $\varphi(i/n)$ равно нулю или почти нулю. Если число n присяжных очень велико, вероятность χ_i вины осуждённого весьма близка единице, какова бы ни была разность $(n - i) - i$ (§ 129). Если, к примеру, подсудимого осудили 520 присяжных против 480, его вину следует считать почти достоверной, хоть её и отрицали 480 присяжных, шанс безошибочности которых, как можно полагать, не отличался от того же шанса у 520 остальных из них. Этого заключения достаточно, чтобы отклонить гипотезу, на которой оно было основано. Действительно, очевидно, что никто не припишет значительное доверие подобному суждению, тем менее такое же доверие, как к почти единогласному решению 1000 присяжных. В соответствии с этой гипотезой, если способности лиц, включённых в общий список, изменяются; если они разнятся в различных районах или в различных видах дел, то вероятность тех шансов не ошибаться, которые более других близки к единице или менее других отличаются от половины, будут повышаться в том же соотношении. Но подобного в действительности не происходит; когда по какой-нибудь причине эта способность повышается, вероятность наиболее близких к достоверности шансов не ошибаться повышается, и противное происходит с шансами, наиболее удалённых от единицы. Принимая функцию φu, которая может удовлетворить этим условиям, и которая кроме того не должна быть нулём или почти нулём при $u < 1/2$, можно будет преодолеть указанные нами трудности. Но для определения этой функции упомянутого недостаточно: этим условиям удовлетворяет бесконечное множество непрерывных и разрывных видов φu и при одном и том же числе n присяжных и одной и той же разности между $(n - i)$ и i они приводят к весьма неравным значениям вероятности χ_i, выраженной формулой (126.1).

Итак, зная эти числа для одного-единственного осуждающего решения, и предполагая, что предварительная вероятность k равна 1/2 или любой другой дроби, нам не удастся, как уже было сказано, определить действительную вероятность верности этого решения. Она зависит от не известного нам шанса не ошибаться, присущего каждому присяжному, и кроме того следует считать невозможным вычисление значения этой вероятности для того, кто только знает, что n присяжных отобраны случайно из общего списка и имеет основание считать, что верность решения зависит лишь от единого для всех n присяжных (§ 122) среднего шанса не ошибаться в этом списке. Действительно, для подобного вычисления придётся сформулировать специальную гипотезу о законе вероятностей значений среднего шанса при его изменении от 0 до 1. Она не должна быть ни гипотезой Лапласа, ни какой-либо иной, недостаточно обоснованной.

Если присяжные, отобранные из общего списка, принимают только одно решение, то предыдущие формулы бесполезны. То же самое произойдёт, если решений немного, но мы знаем, что, напротив, очень большое число осуждений и оправданий в известных соотношениях производится присяжными, последовательно и случайно отобранными из одного и того же общего списка. И на этом соображении основаны, как мы покажем, формулы (117.1, 117.2; 118.1; 119.1, 119.2; 120.1, 120.2), которые содержат лишь две неизвестные постоянные, k и u, и которые поэтому требуют лишь данных двух наблюдений. Определением этих данных мы займёмся в первую очередь.

134. Общий список граждан, которые могут стать присяжными, содержит некоторое число лиц. Каждое жюри состоит из n присяжных, случайно отобранных из этого списка, на один год или на много лет, и они судят очень большое число μ подсудимых. Обозначим через a_i число этих подсудимых, осуждённых большинством по меньшей мере в $(n - i)$ голосов против i. Это предполагает, что i равно нулю или положительному числу, меньшему $n/2$. Предварительный шанс такого осуждения должен изменяться от одного судебного решения к другому, но при любом изменении среднее из его неизвестных значений на протяжении μ решений весьма вероятно будет почти равно отношению a_i/μ (§ 95).

Более того, значения этого среднего шанса и этого отношения очень мало изменяются с числом μ, которое предположено очень большим. И если это число всё более и более возрастает, они будут неопределённо стремиться к специальной константе и достигнут её, если μ может стать бесконечным без какого-либо изменения различных причин осуждения указанным большинством. Эта специальная константа, которую я обозначу через R_i, является суммой шансов, придаваемыми всеми возможными причинами осуждению или рассматриваемому событию, умноженных на соответствующие вероятности этих самых причин (§ 104).

Пронумеровать эти причины и заранее вычислить их влияние невозможно, однако нам нет нужды знать их. Нам достаточно предположить, что ни их вероятности, ни шансы, которые они придают осуждениям, не изменяются, и само наблюдение даст нам знать, соответствует ли это предположение истине. Если соответствует, то, обозначив через a'_i число осуждений большинством не меньшим, чем в $(n - i)$ голосов против i при другом очень большом числе μ′ подсудимых, разность $a'_i/\mu' - a_i/\mu$ весьма вероятно окажется очень небольшой дробью (§ 109). Если же, напротив, эта разность не очень мала, можно будет обоснованно считать, что в промежутке между двумя сериями судебных решений произошло какое-то существенное изменение причин осуждений. Вычисление может лишь уведомить нас о существовании такого изменения, но ничего не сообщит о его сути.

То, что мы говорим об осуждениях большинством по меньшей мере в $(n - i)$ голосов против i, равно относится к осуждениям большинством в $(n - i)$ голосов против i. Обозначим через b_i число последних при μ подсудимых. Существует специальная константа r_i, к которой отношение b_i/μ неопределённо стремится по мере возрастания μ и достигнет её, если μ сможет стать бесконечным без какого-либо изменения причин осуждения. И если b'_i – число осуждений при μ′ подсудимых, разность $b'_i/\mu - b_i/\mu$ весьма вероятно будет очень небольшой дробью. Будут, очевидно, соблюдаться следующие равенства

$$a_i = b_i + b_{i-1} + b_{i-2} + \ldots + b_0,$$
$$a'_i = b'_i + b'_{i-1} + b'_{i-2} + \ldots + b'_0,$$
$$R_i = r_i + r_{i-1} + r_{i-2} + \ldots + r_0.$$

Пусть α будет положительной величиной, очень малой сравнительно с $\sqrt{\mu}$ и $\sqrt{\mu'}$ и

$$P = 1 - \frac{2}{\sqrt{\pi}}\int_\alpha^\infty \exp(-x^2)dx.$$

В соответствии с формулами § 112 величина P окажется также вероятностью того, что границами разностей $a'_i/\mu' - a_i/\mu$ и $b'_i/\mu' - b_i/\mu$ для неизвестных R_i и r_i являются

$$\frac{a_i}{\mu} \mp \alpha\sqrt{\frac{2a_i(\mu - a_i)}{\mu^3}},\ \frac{b_i}{\mu} \mp \alpha\sqrt{\frac{2b_i(\mu - b_i)}{\mu^3}} \qquad (134.1a,\ b)$$

и снова для них

$$\mp\alpha\sqrt{\frac{2a_i(\mu - a_i)}{\mu^3} + \frac{2a'_i(\mu' - a'_i)}{\mu'^3}}, \qquad (134.1c)$$

$$\mp\alpha\sqrt{\frac{2b_i(\mu - b_i)}{\mu^3} + \frac{2b'_i(\mu' - b'_i)}{\mu'^3}}. \qquad (134.1d)$$

При прочих равных условиях протяжённость этих границ укорачивается по мере возрастания μ и μ′ почти обратно пропорционально корням квадратным из этих больших чисел, потому что a_i и b_i возрастают почти как μ, а a'_i и b'_i – почти как μ′. Эта протяжённость также окажется настолько же укороченной, насколько уменьшается α, но их вероятность P понизится вместе с уменьшением α.

135. Как я указал в § 7 Предисловия, все числовые данные, которые я приведу, выбраны из публикуемых правительством *Comptes généraux de l'Administration de la justice criminelle*[9].

С 1825 года до 1830-го включительно количество дел, ежегодно передаваемых на рассмотрение присяжных по всей Франции, составило

5121, 5301, 5287, 5721, 5506, 5068,

а число подсудимых в этих уголовных делах оказалось равным

6652, 6988, 6929, 7396, 7373, 6962,

что ежегодно составляло почти 7 подсудимых на 5 дел. В те же годы число осуждённых большинством по меньшей мере в 7 голосов против пяти составило

4037, 4348, 4236, 4551, 4475, 4130,

так что доли осуждаемых были равны

0.6068, 0.6222, 0.6113, 0.6153, 0.6069, 0.5932.

Уже заметно, что в течение указанных шести лет с неизменным уголовным законодательством эти годичные отношения изменялись очень немного.

За μ я приму общее количество подсудимых за 6 лет, а за a_5 – соответствующее число осуждений: $\mu = 42\,300$, $a_5 = 25\,777$. Соответствующие границы (134.1a) оказываются равными $0.6094 \pm 0.00335\alpha$; если принять, к примеру, $\alpha = 2$, то $P = 0.9953$ окажется очень близкой к достоверности вероятностью того, что без учёта знака неизвестное R_5 отличается от 0.6094 только лишь на 0.0067.

Если разделить рассматриваемый период на две равные части, то для них

$$\mu = 20\,569,\ \mu' = 21\,731,\ a_5 = 12\,621,\ a'_5 = 13\,156,$$
$$a_5/\mu = 0.6136,\ a'_5/\mu' = 0.6054,\ a_5/\mu - a'_5/\mu' = 0.0082.$$

Границами (134.1c) полученной разности будут $\pm\, 0.00671\alpha$; полагая $\alpha = 1.2$, мы получим границы $\pm\, 0.00805$ и $P = 0.9103$,

$1 - P = 0.0897$. Поэтому можно ставить почти 10 против одного за то, что эта разность окажется в пределах ± 0.00805. Хотя фактическая разность ± 0.0082 немного выходит за выведенные границы, и уклонения, и вероятность того, что этого не должно было быть, не настолько значительны, чтобы можно было достаточно обоснованно считать, что в соответствующих причинах произошло какое-нибудь заметное изменение.

В 1831 г. число подсудимых увеличилось до 7606, а число осуждённых дошло до 4098. Закон потребовал, чтобы большинство голосов для осуждения было не меньшим, чем 8 против четырёх. И при этом большинстве оказалось, что

$$\mu = 7606,\ a_4 = 4098,\ a_4/\mu = 0.5388.$$

Если исключить действие нового закона, и если остальные причины, влияющие на решение присяжных, оставались в том году такими же, как и раньше, то отношение b_5/μ можно будет получить, вычитая a_4/μ из a_5/μ, или, иначе, вычитая 0.5388 из найденной выше дроби 0.6094, что приведёт к $b_5/\mu = 0.0706$.

Чтобы проверить этот результат, я замечу, что в 1825 – 1830 гг. закон предусматривал вмешательство судей, составлявших ассизы[10], всякий раз, при котором решение присяжных было принято наименьшим большинством в 7 голосов против пяти. И в *Comptes généraux* сообщается, что в течение последних пяти из этих шести лет количество подобных вмешательств

398, 375, 373, 395, 372

почти не менялось. Всего оно произошло 1911 раз, но соответствующее число подсудимых не было указано. Поэтому следует сравнивать за те годы количество дел, а не подсудимых. Их общее число достигло 26 883, так что [количество дел] $\mu = 26\,883$, $b_5 = 1911$ и $b_5/\mu = 0.0711$, что очень мало отличается от предыдущего результата.

Это соответствие двух значений b_5/μ доказывает, что в 1831 г. вероятности u и k, от которых зависит указанное отношение, оставались почти теми же, что и в предыдущие годы. Следует всё-таки заметить, что вычисление последнего значения было основано на гипотезе о том, что число осуждений семью голосами против пяти так относится к числу подсудимых, как число дел, при котором имело место это большинство, к общему числу дел.

Это нельзя подтвердить априорно, потому что в *Comptes généraux* нет нужных данных.

В 1832 и 1833 гг. количество подсудимых за вычетом их числа в политических процессах составило 7555 и 6964. Существенное различие между этими числами произошло ввиду нового законодательного распоряжения, в соответствии с которым в 1833 г. многие виды дел были переданы из ассизов в ведение полицейских судов. Число осуждений, вынесенных, как и в 1831 г., большинством по меньшей мере в 8 голосов против четырёх, повысилось до 4105 и 4448, так что $a_4/\mu = 0.5887$ и 0.5895.

Видно, что эти отношения очень мало отличаются одно от другого, но их среднее, 05888, превышает $a_4/\mu = 0.5388$ за 1831 г. на 0.05 или примерно на 1/10 этого значения. Если в причинах, которые могли повлиять на голосование присяжных, не произошло никаких изменений, то с учётом границ (134.1c) и их вероятности P это было бы совершенно неправдоподобно. И действительно, с 1832 г. уголовное законодательство подверглось такому изменению: присяжным было предложено учитывать *смягчающие обстоятельства*. Они влекут за собой смягчение наказания, и осуждения стали менее затруднительными и многочисленнее.

136. Различные отношения, которые мы только что привели для всей Франции, не являются неизменными во всех частях королевства. Но если исключить департамент Сена и несколько других, количество уголовных дел, рассмотренных за несколько лет, оказалось не столь большим, чтобы с достаточной вероятностью вывести для каждого ассиза ту постоянную величину, к которой должна стремиться доля осуждённых. Вот результаты для ассиза Парижа.

В 1825 – 1830 гг. число подсудимых ежегодно составляло

802, 824, 675, 868, 908, 804,

а число осужденных

567, 527, 436, 559, 604, 484,

так что доли осуждений оказалась равны

0.7070, 0.6396, 0.6459, 0.6440, 0.6652, 0.6020.

Принимая μ и a_5 равными суммам первых и последних шести чисел, мы получаем $\mu = 4881$, $a_5 = 3177$, $a_5/\mu = 0.6509$. Для Франции в целом первые два числа за те же годы были равны 42 300 и 25 779, и мы установили, что найденное нами соотношение должно очень мало отличаться от 0.6094 [= 25 779:42 300]. Это последнее меньше предыдущего [0.6509] на 0.0416 или примерно на 1/15 своего значения. Однако, с учётом границ (134.1с) и их вероятности P это совершенно неправдоподобно, если только в департаменте Сена какая-то частная причина не привела к менее затруднительным осуждениям, чем в остальной Франции. Какая же это причина? Исчисление не может указать нам её.

Мы всё-таки заметим, что в этом департаменте с населением, едва составляющим 1/36 часть всего населения королевства, число осуждений превышает 1/9 часть осуждений во всей Франции за тот же промежуток времени, т. е. превышает вчетверо. Это обстоятельство делает там подавление преступности более необходимым, и может быть по этой причине тамошние присяжные более суровы[11].

При указанных значениях μ и a_5 границы (134.1a) становятся равными $0.6509 \pm 0.00965\alpha$ и, при $\alpha = 2$, $P = 0.99532$, $1 - P = 0.00468$. Можно ставить более, чем 200 против одного за то, что неизвестное R_5 отличается от 0.6509 лишь на 0.0193 в ту или другую сторону.

Последнее из шести отношений, приведенных выше, 0.6020, заметно меньше среднего из пяти остальных. Имеет смысл исследовать, не в достаточной ли мере эта разность указывает на существование некоторой частной причины, которая могла бы в 1830 г. привести к меньшей суровости присяжных, чем в предыдущие годы. Между тем, принимая для μ и a_5 суммы количеств подсудимых и осуждённых в департаменте Сена за 1825 – 1829 гг., а за μ' и a'_5 соответствующие числа за 1830 г., мы будем иметь

$$\mu = 4077,\ a_5 = 2693,\ \mu' = 804,\ a'_5 = 484,$$
$$a_5/\mu = 0.6605,\ a'_5/\mu' = 0.6019,\ a_5/\mu - a'_5/\mu' = 0.0585.$$

Границы (134.1с) принимают вид $\pm\, 0.02657\alpha$; полагая, что $\alpha = 2$, можно ставить более 200 против одного, за то, что разность $a_5/\mu - a'_5/\mu'$ не должна превышать 0.05314. Фактически же она превосходит эту дробь почти на 1/10 своего значения, и можно полагать, что в то время в голосовании присяжных произошло что-то действительно необычное. Причиной, которая немного поубавила их суровость, могла быть революция 1830 г. и, какова бы она ни была, видимо она действовала на присяжных по всей Франции, потому что в тот год доля осуждений во всём королевстве снизилась почти до 0.59 при своём среднем значении в 0.61 за предыдущие 5 лет.

За 1826 – 1830 гг. включительно число уголовных дел в департаменте Сена возросло до 2963. Из этого числа в 194 случаях осуждение произошло семью голосами против пяти, так что должен был вмешаться суд [ассизов]. Приняв отношение 194/2963 за b_5/μ, мы получим $b_5/\mu = 0.0655$, т. е. немного меньше, чем для Франции в целом.

137. Если, как в *Comptes généraux*, рассматривать по отдельности все виды преступлений, которыми занимались ассизы, количества подсудимых и осуждённых для каждого из этих видов не будут достаточно большими для того, чтобы отношения стали неизменными и могли служить основанием наших вычислений. Но в этих *Comptes* все уголовные дела также разбиты на два класса, *преступлений против личности* и *против собственности*. Эти две большие категории ежегодно приводят к весьма различным отношениям, которые, однако, по отдельности почти неизменны. Их мы и приведём.

За шестилетний период 1825 – 1830 гг. количества подсудимых за преступления в этих двух категориях во всей Франции составляли

1897, 1907, 1911, 1844, 1791, 1666;
4755, 5081, 5018, 5552, 5582, 5296.

Соответствующие числа осуждений при одном и том же уголовном законодательстве оказались равными

882, 967, 948, 871, 834, 766;
3155, 3381, 3288, 3680, 3641, 3364,

а доли осуждений таковыми:

0.4649, 0.5071, 0.4961, 0.4723, 0.4657, 0.4598;
0.6635, 0.6654, 0.6552, 0.6628, 0.6523, 0.6352.

Видно, что в каждой категории эти доли мало изменялись от года к году, но что вторые заметно превышали первые. Принимая за μ и μ′ и a_5 и a'_5 суммы, соответственно, подсудимых и осуждений для преступлений соответственно против личности и против собственности, мы получим

$$\mu = 11\,016,\ a_5 = 5268,\ \mu' = 31\,284,\ a'_5 = 20\,509,$$
$$a_5/\mu = 0.4782,\ a'_5/\mu' = 0.6556.$$

Второе отношение превышает первое немного больше, чем на треть своего значения. Исходя из этих чисел, можно заключить, что границами (134.1а) неизвестного R_5 в указанных случаях являются

$$0.4782 \pm 0.00675\alpha \text{ и } 0.6556 \pm 0.00380\alpha.$$

Следовательно, при $\alpha = 2$ с вероятностью, очень близкой к достоверности, в рассматриваемых случаях это неизвестное R_5 не отличается от 0.4782 более, чем на 0.0135, и от 0.6556 более, чем на 0.0076.

В 1831 г. для осуждений потребовалось большинство не менее, чем в 8 голосов против четырёх. Оказалось, что, соответственно,

$$\mu = 2046,\ a_4 = 743,\ \mu' = 5560,\ a'_4 = 3355,$$
$$a_4/\mu = 0.3631,\ a'_4/\mu' = 0.6034.$$

При вычитании полученных отношений из предыдущих оказывается, что доли осуждённых прежним наименьшим большинством в 7 голосов против пяти были равны $b_5/\mu = 0.1151$ и $b'_5/\mu' = 0.0522$. Примечательно, что первое из них, относящееся к преступлениям против личности, почти вдвое превышает второе, тогда как, напротив, a'_5/μ' примерно на треть превышает a_5/μ.

Таким образом, осуждения при преступлениях против собственности не только пропорционально многочисленнее, но и выносятся более значительным большинством.

Рассматриваемые отношения совсем не одинаковы для обоих полов. Ежегодно женщины почти постоянно составляли 0.18 от всех подсудимых в ассизах. За пятилетний период 1826 – 1830 гг. количества подсудимых женщин, обвиняемых в преступлениях тех же двух категорий, μ и μ′, и количества осуждённых из них, a_5 и a'_5, были равны

$$\mu = 1305,\ \mu' = 5465,\ a_5 = 586,\ a'_5 = 3312,$$

$a_5/\mu = 0.4490$, $a'_5/\mu' = 0.6061$.

При сравнении этих отношений с предыдущими a_5/μ и a'_5/μ' оказывается, что они меньше, но лишь примерно на 1/16 и 1/12 своих значений.

В 1832 и 1833 гг., когда осуждения выносились большинством не меньшим, чем в 8 голосов против четырёх, притом с учётом *смягчающих обстоятельств*, подсудимых и осуждённых мужчин и женщин было

$\mu = 4108$, $\mu' = 10\,421$, $a_4 = 1889$, $a'_4 = 6664$,
$a_4/\mu = 0.4598$, $a'_4/\mu' = 0.6395$.

Как и выше, штрихованные буквы относятся к преступлениям против собственности. При $\alpha = 2$ границы (134.1а) указывают, что с вероятностью, очень близкой к достоверности, неизвестное R_5 отклоняется от 0.4598 не более, чем на 0.022 и от 0.6395 не более, чем на 0.0133 соответственно. Можно заметить, что отношение a_4/μ к a'_4/μ' оказалось почти равным найденному выше отношению a_5/μ к a'_5/μ'. Сравнивая a_4/μ и a'_4/μ' с аналогичными значениями 1831 г., можно также указать, что влияние *смягчающих обстоятельств* увеличило отношение a'_4/μ' для преступлений против собственности всего на 1/15, второе же отношение для преступлений против личности, a_4/μ, почти на треть своего значения.

138. В § 122 было установлено, что шанс подсудимого быть осуждённым присяжными, случайно отобранными из общего списка департамента или ассиза, оказался бы тем же самым при едином шансе не ошибаться у всех из них. При осуждении большинством не меньшим, чем в $(n - i)$ голосов против i, шанс осуждения поэтому выражается первой из формул (118.1), а при большинстве в $(n - i)$ голосов против i – формулой (117.1). И для каждого департамента и каждой категории уголовных дел величины c_i и γ_i, выраженные этими формулами, оказываются теми, к которым отношения a_i/μ и b_i/μ неопределённо стремятся с дальнейшим возрастанием весьма большого по предположению μ. Иначе говоря, если рассматривать дела той же самой категории, в том же самом департаменте и отдельно для обвиняемых мужчин и женщин, c_i и γ_i совпадут с неизвестными R_i и r_i (§ 134).

Все виды уголовных дел мы будем, как и выше, подразделять на две чёткие категории, преступлений против личности и против собственности, и штрихованные буквы будут относиться к этой второй категории. Однако, чтобы не усложнять вычисления слишком сильно, мы не будем учитывать пол подсудимых, влиянием которого на долю осуждений можно пренебречь, имея в виду, что из общего числа подсудимых более 5/6 составляли мужчины. Для каждого департамента мы получим

$$a_i/\mu = c_i,\ b_i/\mu = \gamma_i,\ a'_i/\mu' = c'_i,\ b'_i/\mu' = \gamma'_i, \qquad (138.1a, b, c, d)$$

с тем лучшим приближением и тем более высокой вероятностью, чем больше были числа μ и μ'.

Если отношения в левых частях уравнений (138.1) известны для различных департаментов, этих уравнений будет достаточно, чтобы определить неизвестные k и u, содержащиеся в величинах c_i и γ_i и их аналогов k' и u' в c'_i и γ'_i. Но в настоящее время необходимость иметь значительные μ и μ' исключает возможность применения уравнений (138.1) к каждому департаменту по отдельности. Приходится предполагать, что неизвестные u, u', k, k' как правило мало изменяются от одного из них к другому, так что для указанных левых частей можно будет принимать данные, относящиеся ко всей Франции.

Величины u и u', определённые таким образом, в точности совпадут с шансами не ошибаться, которые будут иметь место, если списки присяжных всех департаментов объединить в единый список и случайно отбирать из него каждого присяжного. При этой гипотезе величины k и k', поскольку они зависят от умения магистратов, которые руководят предварительным следствием, могут, тем не менее, быть различными в различных департаментах. Однако, по отношению к этим неизвестным уравнения (138.1) линейны, и их выводимые значения являются средними из тех, которые в действительности относятся ко всем департаментам. В заключение следует заметить, что необходимость удовлетворяться этими общими значениями u, u', k, k' происходит исключительно ввиду отсутствия полных данных, а не какого-то несовершенства излагаемой нами теории.

Величины c_i и γ_i не изменятся при замене k и u на $(1 - k)$ и $(1 - u)$ (§§ 117 и 118). Поэтому, если данным a_i/μ и b_i/μ соответствуют k и u, большие, чем 1/2, удовлетворяющие уравнениям (138.1a, b) то есть и другие k и u, также удовлетворяющие этим уравнениям, но меньшие, чем 1/2. Между тем, мы должны предположить, что предварительная средняя вероятность вины подсудимого превышает вероятность его невиновности и что средний шанс присяжного не ошибаться превышает 1/2. Таким образом, следует применять значения k и u, большие, чем 1/2 и отклонять остальные, как посторонние для нас.

То же замечание относится к уравнениям (138.1c, d) и к значениям выводимых из них величин k' и μ'. Впрочем, если применять эти уравнения к многочисленным политическим процессам, происшедшим в злополучные годы революции, можно будет, как описано в § 12 Предисловия, применять их корни, меньшие, чем 1/2, потому что в то время предварительная правовая невиновность подсудимых могла быть вероятнее их вины, а вероятность умышленной ошибки присяжных могла превышать их шанс не ошибаться.

139. Я принимаю в формулах (117.1) и (118.1) $n = 12$ и $i = 5$. Их коэффициенты будут такими:

$$N_0 = 1,\ N_1 = 12,\ N_2 = 66,\ N_3 = 220,\ N_4 = 495,\ N_5 = 792.$$

Я также полагаю, что

$$\frac{a_5}{\mu} = c,\ \frac{b_5}{\mu} = 792\gamma,\ u = \frac{t}{1+t},\ 1-u = \frac{1}{1+t}$$

и уравнение (138.1b) примет вид

$$\gamma = \frac{(t^2+1)t^5}{2(1+t)^{12}} + \frac{(2k-1)(t^2-1)t^5}{2(1+t)^{12}}. \tag{139.1}$$

Замечая, что

$$U_5 = 1 - 924u^6(1-u)^6 - V^5,$$

уравнение (138.1) можно записать в виде

$$c = k[1 - \frac{924t^6}{(1+t)^{12}}] - \frac{(2k-1)}{(1+t)^{12}}[1 + 12t + 66t^2 + 220t^3 + 495t^4 + 792t^5]. \tag{139.2}$$

Уравнения (139.1) и (139.2) соответствуют преступлениям против личности. Соответствующие уравнения для второй категории преступлений устанавливаются заменой в них величин c, γ, k, t на c', γ', k', t'. Неизвестное t может принимать все значения от $t = 0$ до ∞, при значениях $u = 0$ и 1. Но значения $t > 1$ соответствуют $u > 1/2$, и их-то и следует допустить. Далее, неизвестное k должно быть заключено в границах 1/2 и 1, поэтому, см. уравнение (139.1), t должно быть таким, чтобы

$$\gamma > \frac{(t^2+1)t^5}{2(1+t)^{12}},\ \gamma < \frac{t^7}{(1+t)^{12}}. \qquad (139.3a, b)$$

Соответственно должны быть определены его границы. Заметим по этому поводу, что функция от t в первом неравенстве постоянно убывает от $t = 0$ до ∞, а вторая функция вначале возрастает от $t = 1$ до 7/5, затем убывает до $t = \infty$.

Исключая k из уравнений (139.1) и (139.2), можно вывести возвратное уравнение 24-й степени относительно t, которое поэтому может быть сведено к уравнению 12-й степени, но намного проще вычислить одновременно k и u, удовлетворяющие уравнениям (139.1) и (139.2), последовательными приближениями.

140. Для шестилетнего периода 1825 – 1830 гг. мы имеем

$c = 0.4782$, $\gamma = 0.1151/792 = 0.0001453$.

При $t = 2$ дробь в неравенстве (139.3a) превысит это значение γ, а при $t = 3$ γ превысит вторую функцию от t. Поэтому значение t должно быть больше двух и меньше трёх. Легко удостовериться, что в указанных границах это неизвестное имеет лишь одно возможное значение. После некоторых попыток я принял это значение равным 2.112, и из уравнения (139.1) тогда последовало значение $k = 0.5354$. Подставляя эти значения в правую часть уравнения (139.2), мы находим, что она равна 0.4783, что отличается от левой части лишь на 0.0001. Поэтому с очень хорошим приближением

$k = 0.5354$, $t = 2.112$.

Для тех же лет

$c' = 0.6556$, $\gamma' = 0.0523/792 = 0.00006604$.

Я подставляю эти значения вместо c и γ в уравнения (139.1) и (139.2) и заменяю t и k на t' и k'. Решая их как в предыдущем случае, я нахожу с той же степенью приближения, что

$$k' = 0.6744,\ t' = 3.4865,$$

так что

$$u = \frac{t}{1+t} = 0.6786,\ u' = \frac{t'}{1+t'} = 0.7771$$

оказываются для тех лет шансами какого-либо присяжного не ошибаться в делах о преступлениях против личности и против собственности.

До объявления решения по делу кто-либо, не знающий ни присяжных, ни даже местоположения суда, мог бы в те годы ставить немногим более двух против одного за то, что ни один присяжный не ошибается при голосовании в делах первой категории, и почти семь против двух за то же самое в делах второй категории. В таких случаях применяют обычное выражение *ставить что-то против чего-то*, чтобы лучше прояснить значимость, которую следует придавать u и u', а также потому, что предположенное пари является обманчивым, поскольку никогда нельзя будет выяснить, кто выиграл. То лицо, которое ничего не знало об обстоятельствах дела, могло держать пари, основываясь и на предшествовавших значениях k и k' и ставить немного меньше семи против шести за то, что подсудимый виновен в деле первой категории, и немного больше двух против одного за то же в деле второй категории. В дальнейшем мы увидим, какой оказывается вероятность вины подсудимого после решения его дела.

Если рассматривать количества подсудимых и осуждённых без различия категорий преступлений, потребуется принять, снова для тех же лет и для Франции в целом,

$$c = 0.6094,\ \gamma = 0.0706/792 = 0.00008914.$$

Решая уравнения (139.1) и (139.2), мы тогда установим, что

$$k = 0.6391,\ t = 2.99,\ u = 0.7494.$$

Если же отдельно рассматривать департамент Сена, надо будет применить значения c и γ (§ 136)

$$c = 0.6509,\ \gamma = 0.0655/792 = 0.00008267,$$

которые приводят к

$$k = 0.678,\ t = 3.168,\ u = 0.7778.$$

В те же годы, не учитывая категории преступлений, вероятности k и u в парижском ассизе оказались немного выше, чем в остальном королевстве и немного превысили 2/3 и 3/4. Однако, разности двух значений каждой из этих величин незначительны, что является одной из причин полагать, что значение каждой из них в одной части Франции то же, что и в другой части. Это обосновывает насколько возможно гипотезу об их равенстве по всему королевству, которую мы и приняли, чтобы можно было вычислить их приближённые значения, исходя их достаточно большого числа наблюдений.

Таким образом, как было сказано раньше, в 1831 г. значения k и u или k' и u' оставались неизменными, но они должны были измениться в последующие годы вместе с отношениями, которые выводятся по ним. Мы знаем отношения a_4/μ и a'_4/μ' только для 1832 и 1833 гг., чего недостаточно для определения неизвестных u и k или u' и k'. Заметим также, что эти величины быть может изменились вторично после последнего закона, который предписал тайное голосование, сохранив требование учёта *смягчающих обстоятельств*, что могло повлиять на их [!] шанс не ошибаться.

Мы не можем поэтому определить значения u и k или u' и k' после 1831 г. Но этот закон, установив, что для осуждения достаточно большинства в 7 голосов против пяти, потребовал, чтобы присяжные сообщали, не приняли ли они своё решение минимальным большинством. Если впредь *Comptes généraux* будут указывать число осуждённых, а не только количество дел, по которым решение было принято этим наименьшим большинством; и если станет известно, сколько было мужчин и женщин среди подсудимых, и сколько из них проходило по каждой категории преступлений, – если эти условия будут выполнены, то через несколько лет окажется возможным очень точно определять u и k для различных частей королевства, для мужчин и женщин, и для этих категорий преступлений.

141. Зная значения u и k, можно из формул (117.1, 117.2, 118.1) определить вероятности осуждения и оправдания при заданном большинстве голосов или при большинстве голосов, не меньшим заданного. При $n = 12$ и $i = 0$

$$\gamma_0 = ku^{12} + (1-k)(1-u)^{12},\ \delta_0 = (1-k)u^{12} + k(1-u)^{12}$$

окажутся этими вероятностями единогласного осуждения или оправдания, и потому

$$\gamma_0 + \delta_0 = u^{12} + (1-u)^{12}$$

и будет указанной вероятностью. Кроме того, величина

$$\gamma_0 - \delta_0 = (2k-1)[u^{12} - (1-u)^{12}]$$

положительна ввиду условий $k > 1/2$ и $u > 1 - u$, так что единогласное решение дела менее вероятно при оправдании, чем при осуждении. Если шанс u не ошибаться существенно отличается от 0 и 1, эти различные вероятности оказываются весьма низкими. Приняв, к примеру, значения u и k для всей Франции без различия категорий преступлений, т. е. полагая, что $k = 0.6391$ и $u = 0.7494$, мы получим

$$\gamma_0 = 0.021,\ \delta_0 = 0.013,\ \gamma_0 + \delta_0 = 0.0314.$$

Этого достаточно, чтобы доказать, насколько редкими должны быть единогласные решения 12 присяжных. Если же и для осуждения, и для оправдания требуется единогласие, то в соответствии со значением $\gamma_0 + \delta_0$ можно ставить почти 32 против одного за то, что никакого решения не будет принято. И если присяжные не общаются друг с другом и не соглашаются принять решение простым большинством, то так и будет происходить примерно в 32 делах из 33.

Обозначим через M вероятность того, что одно из μ дел было решено либо единогласно, либо нет, тогда

$$M = (1 - \gamma_0 - \delta_0)^{\mu},$$

и если желательно, чтобы $M = 1/2$, то, при прежнем значении суммы $\gamma_0 + \delta_0$

$$\mu = \frac{-\lg 2}{\lg(1-\gamma_0-\delta_0)} = 21.73.$$

Это означает, что можно ставить немного больше одного против одного за то, что по меньшей мере решение одного дела из 22 будет единогласным. И невыгодным окажется подобное пари для 21 дела.

142. Прежде, чем идти дальше, необходимо напомнить то, что было сказано в § 2 Предисловия о смысле, который мы придаём в решении дела слову *виновен,* и вывести несколько важных следствий.

Объявляя, что подсудимый виновен, присяжный утверждает, что для осуждения собрано, по его мнению, достаточно улик. Под решением, что подсудимый невиновен, поэтому разумеют, что для осуждения вероятность вины недостаточно высока. Этот последний случай не означает, что присяжный считает подсудимого невиновным, более часто он несомненно полагает, что подсудимый, пожалуй, виновен. Иногда вероятность вины подсудимого превышает 1/2, оставаясь, однако, ниже того значения, при котором и совесть присяжного, и общественная безопасность потребовали бы осуждения. Реальный смысл того или иного решения присяжного состоит в том, что подсудимый *подлежит или не подлежит осуждению*. Поэтому вероятности P_i и Q_i верности осуждения или оправдания (§ 120) выражают и наши основания верить, что осуждённый подлежал осуждению, а оправданный не подлежал.

Таким образом, P_i несомненно ниже действительной вероятности вины осуждённого, и, напротив, Q_i превосходит вероятность невиновности оправданного. Эти другие вероятности никак не могут быть вычислены, что, однако, не относится к вероятностям P_i и Q_i, определяемым таким образом и рассматриваемым при очень большом числе решений по делам одного и того же вида. Не следует также полагать, что P_i и Q_i выражают общее мнение; они выражают вероятности осуждения или оправдания судом, состоящим из всех граждан, включённых в общий список, из которого по жребию отбираются 12 присяжных.

Шанс c_i осуждения каким-то числом присяжных ниже дроби, которую мы обозначили через k (§ 118) и которая, как правило, намного ниже P_i. Аналогично, шанс d_i оправдания всегда ниже дроби $(1 - k)$, которая в свою очередь намного ниже Q_i. Для присяжных каждого ассиза и для каждой из двух категорий преступлений следует поэтому представлять, что существует определённая вероятность z, которая считается достаточной и необходимой для осуждения. И шанс u не ошибаться для присяжного, отобранного по жребию из списка своего департамента, равен вероятности, с которой он решает, является ли виновность подсудимого равной или превосходящей z или нет. Вероятность u в основном зависит от степени образованности класса лиц, включённых в общий список присяжных и вероятности z, т. е. сложившегося у них мнения о необходимости более или менее сильного подавления различных видов преступности. Эти две различные вероятности могут, следовательно, изменяться во времени и от одного департамента к другому. Понятно, как u может быть выведено из наблюдений, но у нас нет никакой возможности установить z. Мы можем только заключить, что при прочих равных условиях оно возрастает и убывает, если заметим, что доля осуждённых заметно снизилась или повысилась. Так, когда вопрос о *смягчающих обстоятельствах* был поставлен перед присяжными, и оказалось, что эта доля возросла от 0.54 до 0.59 (§ 135), пришлось заключить, что, кроме как при положительных решениях [?], при осуждениях они приняли более низкую, чем прежде вероятность z, потому что наказания стали менее тяжёлыми.

Предварительная вероятность вины подсудимого, несомненно, намного превышает ту, которую мы обозначили через k. Наибольшее найденное нами значение k почти равнялось 3/4, но тем не менее никто не стал бы колебаться поставить намного более трёх против одного за то, что кто-то, представший перед судом ассизов, действительно виновен. Но сказанное о P_i равным образом относится к k; следует также понять, что k выражает лишь существовавшую вероятность того, что подсудимый *подлежит осуждению*. Эта вероятность может поэтому зависеть от той, которую присяжные требуют для осуждения, но которая по своей природе не зависит от вероятности u того, что кто-либо из присяжных не ошибается. Следовательно, k может изменяться вместе с вероятностью z даже если формы предварительного следствия и способности тех, кто руководил им, оставались теми же самыми, притом при любой вероятности u. Вот пример такого изменения.

В период 1814 – 1830 гг. уголовные дела в Бельгии[12] решались трибуналами пяти судей, и для осуждения было достаточно большинства в 3 голоса против двух. В 1830 г. состав трибуналов изменился, и в 1831 г. было восстановлен суд присяжных, существовавший при французском владычестве, и необходимое для осуждения большинство оказалось равным семи голосам против пяти. Формы предварительного следствия остались прежними.

Из *Comptes de l'administration de la justice criminelle* в этом королевстве, недавно опубликованных [бельгийским] правительством, следует, что в 1832 – 1834 гг. доля осуждённых составляла 59, 60 и 61 сотых. Заметно, что она изменялась очень мало и что её среднее почти равно среднему для Франции до 1830 г. Но в этих *Comptes* ничего не сказано о количестве осуждений ни наименьшим большинством в 7 голосов против пяти, ни каким-либо иным определённым большинством, и потому приведенная нами доля [?] недостаточна для вывода значений u и k в Бельгии. Однако, полная доля, т. е. та, которую мы обозначили через a_5/μ, так мало отличается в Бельгии от французских данных, что можно полагать, что и частичное отношение b_5/μ во Франции и Бельгии почти одно и то же, и что, следовательно, значения u и k в этих странах почти те же. Можно поэтому считать, что в Бельгии значение k не отличается намного от дроби 0.64, которая была раньше получена для Франции в целом без различия категорий преступлений.

В тех же *Comptes* мы находим, что в 1826 – 1829 гг. те же почти равные друг другу доли доходили до 84, 85, 83 и 81 сотых, а их среднее немного превышало 0.83. Но, в соответствии с указанным в § 118, вероятность, приближённым значением которой является это среднее, должна неизменно быть ниже, чем k. Следовательно, в те годы k должно было быть намного большим, чем в 1832 – 1833 гг. Это может быть приписано только изменению неизвестного z в эти два периода, т. е. тому, что присяжные требовали для осуждения более высокую вероятность, чем достаточную для судей. Это заключение к тому же независимо от шанса u не ошибаться, который мог измениться или нет, т. е. мог быть выше или ниже у судей, чем у присяжных. Что именно имело место сказать нельзя из-за отсутствия необходимых наблюдений.

Величина k зависит от вероятности z, и поэтому неравенство её значений в двух рассматриваемых нами категориях преступлений могло последовать по двум различным причинам. Презумпцию вины труднее установить в делах о преступлениях против личности, а для осуждения присяжные требуют в этих делах более высокую вероятность z. Можно полагать, что, действуя совместно, эти две различные причины и привели к отмеченному неравенству.

Из зависимости между z и k следует, что если в 1832 и 1833 гг. учёт проблемы *смягчающих обстоятельств* заметно снизил вероятность, которую присяжные считали достаточной для осуждения, вероятность k, напротив, должна была возрасти, и изменение u и k в противоположных направлениях должно было также привести к возрастанию u. Действительно, можно предположить, что шанс присяжных не ошибаться уменьшился, потому что, с одной стороны, они требовали более низкую вероятность для осуждения, и, с другой стороны, до решения дела существовала более высокая вероятность того, что подсудимый подлежал осуждению.

143. Нам по существу осталось при помощи формул (120.1) и (120.2) и найденных значений u и k или u' и k' вычислить вероятности вины осуждённого и невиновности оправданного; или, говоря точнее, вычислить вероятности того, что в первом случае подсудимый подлежал осуждению, а во втором – не подлежал. Но прежде всего мы преобразуем указанные формулы в уравнения, более удобные для вычисления, и добавим другие формулы, численные значения [параметры] которых также очень важно знать.

Ввиду уравнения (118.1а) формулу (120.1) можно заменить уравнением

$$P_i c_i = kU_i,$$

в котором за приближённое значение c_i принимается известное из наблюдений отношение a_i/μ. Величина $(1 - P_i)$ есть вероятность того, что осуждённый большинством не меньшим, чем в $(n - i)$ голосов против i, невиновен; c_i – вероятность того, что подсудимый, будь он виновен ли нет, осуждён этим большинством. Таким образом, произведение $c_i(1 - P_i)$ выражает шанс невиновного подсудимого быть тем не менее осуждённым.

Обозначив его через D_i и учитывая предыдущее уравнение и уравнение (118.1а), мы получим

$$D_i = (1 - k)V_i.$$

Этот результат можно было также вывести, применив рассуждение в § 120, при помощи которого там было получено P_i.

Если число голосов, необходимых для осуждения, не менее

$(n - i)$, то вероятность Π_i того, что оправданный подсудимый невиновен, определяется по Q_i или по формуле (120.2) при замене i на $(n - i - 1)$. Учитывая уравнение (118.1b), мы заключаем, что

$$\Pi_i d_{n-i-1} = (1 - k)U_{n-i-1},$$

или, что то же самое, в соответствии с § 118

$$\Pi_i(1 - c_i) = (1 - k)(1 - V_i).$$

Вероятность того, что оправданный подсудимый виновен, равна $(1 - \Pi_i)$, а вероятность подсудимому не быть осуждённым равна $(1 - c_i)$. Таким образом, произведение $(1 - \Pi_i)(1 - c_i)$ выражает шанс виновного подсудимого быть тем не менее оправданным. Обозначим его через Δ_i, тогда

$$\Delta_i = 1 - c_i - (1 - k)(1 - V_i)$$

или, ввиду уравнения (118.1a),

$$\Delta_i = k(1 - U_i).$$

Шансы D_i и Δ_i являются, так сказать, мерами опасности осудить подсудимого, не подлежащего осуждению, и обществу видеть подлежащего осуждению подсудимого оправданным. По отношению к истинной вине или невиновности подсудимого нельзя забывать, что D_i в отличие от P_i является только верхней границей и что Δ_i в отличие от Q_i является только нижней границей. После вычисления P_i и Π_i немедленно определяются D_i и Δ_i, ибо ввиду предыдущих уравнений

$$D_i = 1 - k - \Pi_i(1 - c_i),\ \Delta_i = k - P_i c_i,$$

так что заметно, что шансы Δ_i и D_i всегда соответственно меньше вероятностей k и $(1 - k)$ предварительных вины и невиновности. При очень большом числе μ подсудимых количества осуждений и оправданий, данные из наблюдений, будут равны a_i и $(\mu - a_i)$. Числа осуждённых невиновных и оправданных виновных будут весьма вероятно почти равны произведениям $D_i a_i$ и $\Delta_i(\mu - a_i)$.

При $n = 12$ и $i = 5$ и 4, принимая a_5/μ и a_4/μ за приближённые значения c_5 и c_4 и подставляя, как и раньше, $t/(1 + t)$ и $1/(1 + t)$ вместо u и $(1 - u)$, мы выведем из предыдущего уравнения

$$\frac{a_5}{\mu}P_5 = k(\Omega + \frac{924t^6}{(1+t)^{12}}),\ \ \frac{a_4}{\mu}P_4 = k(\Omega + \frac{924t^6 + 792t^7}{(1+t)^{12}}),$$

$$[1-\frac{a_5}{\mu}]\Pi_5 = (1-k)\Omega,\ [1-\frac{a_4}{\mu}]\Pi_4 = (1-k)(\Omega - \frac{792t^5}{(1+t)^{12}}),$$

$$\Omega = [1 - \frac{1+12t+66t^2+220t^3+495t^4+792t^5}{(1+t)^{12}}].$$

В то же время

$D_5 = 1 - k - (1 - a_5/\mu)\Pi_5$, $\Delta_5 = k - (a_5/\mu)P_5$,
$D_4 = 1 - k - (1 - a_4/\mu)\Pi_4$, $\Delta_4 = k - (a_4/\mu)P_4$.

Таковы различные формулы, которые [чьи параметры] следовало свести к числам. Величины в них относятся к преступлениям против личности. Аналогичные величины, относящиеся к преступлениям против собственности, мы обозначили теми же буквами со штрихами.

144. В течение 1831 г. большинство, необходимое для осуждения, составляло 8 голосов против четырёх, а проблемы *смягчающих обстоятельств* не существовало. Мы имели

$(a_4/\mu) = 0.3632$, $t = 2.112$, $k = 0.5354$,

откуда следовало, что

$P_4 = 0.9811$, $\Pi_4 = 0.7186$, $D_4 = 0.00689$, $\Delta_4 = 0.1791$.

Из 743 осуждённых в тот год почти 5, как показывает это значение D_4, не должны были быть осуждены, а Δ_4 устанавливает, что примерно 233 подсудимых из 1303 не должны были быть оправданы. Шанс быть осуждённым, хоть и не подлежавшим осуждению, чуть превышал 1/150, а шанс быть оправданным, хоть и подлежавшим осуждению, был заключён между 1/6 и 1/5. Наконец, вероятность вины осуждённого не отличалась на 1/50 от достоверности, а вероятность невиновности оправданного, т. е. недостаточной доказанности вины, ненамного превышала 2/3.

Эти результаты относились к преступлениям против личности. В том же году относительно преступлений против собственности было

$(a'_4/\mu') = 0.6034$, $t' = 3.4865$, $k' = 0.6744$

и потому

$P'_4 = 0.9981$, $\Pi'_4 = 0.8199$, $D'_4 = 0.0004$, $\Delta'_4 = 0.0721$.

Для этой категории преступлений доля ошибочно осужденных составила таким образом лишь 4 на 10 000, т. е. менее двух на 3355 объявленных осуждений. Доля оправданных, подлежавших осуждению, превзошла 7/100, и их число должно было составить примерно 159 из 2205 оправданий. Вероятность вины осуждённого отличалась от достоверности менее, чем на 2/1000, а вероятность невиновности оправданного оказалась чуть выше 4/5. Заметно, что эти результаты удовлетворительнее тех, которые относились к преступлениям против личности, потому что осуждения, хоть относительно и многочисленнее, весьма вероятно принимались более квалифицированным большинством (§ 141).

Восемь вероятностей P_4, P'_4, …, которые мы только что вычислили, основаны на соотношениях a_4/μ, a'_4/μ', b_4/μ, b'_4/μ', данных из наблюдений и использованных раньше для вычисления t, t', k, k'. Все эти восемь величин – дроби, меньшие единицы, что является тем более примечательным подтверждением теории, поскольку при произвольных t и k и t' и k', хоть и ненамного отличных от полученных из наблюдений, этот результат уже не имеет места с той же общностью.

В годы, предшествовавшие 1831 г., достаточным большинством для осуждения было 7 голосов против пяти, однако при минимальном большинстве вмешивался суд [ассизов], и осуждение становилось окончательным, только если большинство из пяти судей, из которых суд состоял в то время, соглашалось с большинством присяжных. По этой причине необходимо рассматривать отдельно осуждения, вынесенные этим наименьшим большинством, и большинством не менее, чем в 8 голосов против четырёх. Во втором случае значениями вероятностей P_4 и P'_4, Π_4 и Π'_4, D_4 и D'_4, Δ_4 и Δ'_4 являются вычисленные нами чуть выше, потому что и до 1831 г., и в том году значения t и k и t' и k' были теми же самыми (§ 137).

Итак, в период 1825 – 1830 гг. примерно 5000 человек было осуждено этим большинством, не меньшим, чем в 8 голосов против четырёх, за преступления против личности и почти 20 000 – за преступления против собственности. И в соответствии с предыдущими значениями D_4 и D'_4, примерно 35 и 8 человек весьма вероятно не подлежали осуждению. Это, несомненно, слишком много, если желательно сказать, что они действительно были невиновны.

По отношению к другим осуждениям, объявленным наименьшим большинством в 7 голосов против пяти, вероятность того, что осуждённые были виновны, если в формуле (119.1) положить $n = 12$, $i = 5$, $u = t/(1 + t)$, $1 - u = 1/(1 + t)$, равна

$$p_5 = \frac{kt^2}{kt^2 + 1 - k}.$$

Для преступлений против личности, как и выше, $t = 2.112$, $k = 0.5354$ и поэтому $p_5 = 0.8372$. Для преступлений против собственности, подставляя p'_5, k', t' вместо p_5, k, t и принимая, как и выше, что $t' = 3.4865$, $k' = 0.6744$, мы получаем $p'_5 = 0.9618$.

Наконец, если не различать этих категорий преступлений, снова для Франции в целом следует принять $k = 0.6391$ и $t = 2.99$ (§ 140). Обозначая через w_5 соответствующее значение p_5 вероятности вины осуждённого, мы имеем $w_5 = 0.9406$.

Вычитая из единицы значения p_5, p'_5, w_5, мы получим почти точно 16, 4 и 6 сотых для вероятности ошибки судебного решения в трёх рассмотренных случаях. По формуле Лапласа (§ 132) она должна быть одной и той же во всех этих случаях и равняться 0.29, т. е. быть почти вдвое выше значения $(1 - p_5)$ и впятеро выше $(1 - w_5)$. В следующем параграфе мы увидим, к чему сводится эта вероятность $(1 - w_5)$ невиновности подсудимого, если его осуждение подтвердил суд ассизов большинством не меньшим, чем в 3 голоса против двух.

При объединении данных, достаточных для вычислений, так, как было сказано в § 140, из значений k и u или k' и u', которые имели место в рассматриваемую эпоху, подобно предыдущему выводятся соответствующие вероятности P_5, Π_5, D_5, Δ_5 или P'_5, Π'_5, D'_5, Δ'_5 и сравниваются с вероятностями P_4, Π_4, D_4, Δ_4 или P'_4, Π'_4, D'_4, Δ'_4, которые мы определили для 1831 г. И после этого мы сможем без всяких иллюзий выяснить, например, относительные преимущества уголовного законодательства в эти две эпохи и для общественной безопасности, и для должных гарантий, предоставляемых подсудимым.

При тех же самых данных наблюдения им, по замечанию в § 138, будут соответствовать две различные пары значений k и u или k' и u', большими 1/2 или меньшими этих значений и дополняющими их до единицы. Мы, к примеру, нашли, что для 1831 г. и преступлений против собственности $k' = 0.6744$, $u' = 0.7771$, но используя те же данные наблюдений мы также определили, что

$$k' = 1 - 0.6744 = 0.3256,\ u' = 1 - 0.7771 = 0.2229.$$

Значение u' изменилось на $(1 - u')$ и в то же время t' изменилось на $1/t'$, так что $t' = 1/3.4865 = 0.2868$. Неизменно принимая $a'_4/\mu' = 0.6034$, мы находим, что $P'_4 = 0.000675$, так что осуждение не повысило вероятность вины подсудимого, а напротив понизило её почти до нуля. Но, как мы уже сказали в § 138, в принципе следует отбрасывать значения неизвестных k и u или k' и u', меньшие значений противоположных вероятностей. Тем не менее, они указываются вычислениями, чтобы включить случай появляющихся при очень большом числе особых судебных решений, при которых законосообразная вина осуждённого менее вероятна, чем его невиновность.

145. Если в формуле (118.1а) принять $n = 5$, $i = 2$, $u = t/(1 + t)$, $1 - u = 1/(1 + t)$, то окажется, что вероятность того, что подсудимый будет осуждён трибуналом пяти судей большинством не меньшим, чем в 3 голоса против двух, равна

$$c_2 = k - \frac{(2k-1)(1+5t+10t^2)}{(1+t)^5}.$$

Здесь, как всегда, k обозначает предварительную вероятность вины этого подсудимого, а u – шанс того, что никто из судей не ошибается. Ввиду формулы (120.1) мы в то же время имеем

$$c_2 P_2 = k[1 - \frac{1+5t+10t^2}{(1+t)^5}]$$

или проще, учитывая предыдущее уравнение,

$$(2k - 1)c_2 P_2 = k(k - 1 + c_2)$$

для определения вероятности P_2 вины подсудимого после его осуждения.

Применим эти уравнения к случаю подсудимого, осуждённого минимальным большинством в 7 голосов против пяти и затем отданного под суд ассизов, как это происходило до 1831 г. За k принимается вероятность вины подсудимого, вытекающая из решения присяжных; приближённое и весьма вероятное значение c_2 устанавливается по наблюдениям и оказывается равным доле осуждений, вынесенных ассизами по очень большому числу дел.

Из *Comptes généraux* видно, что за пятилетний период 1826 – 1830 гг. ассизам королевства было передано 1911 дел после осуждения подсудимых большинством в 7 голосов против пяти. Ассизы подтвердили осуждения в 1597 случаях, но *Comptes* не указывают, как эти числа были распределены по категориям преступлений, и мы вынуждены определять P_2 и неизвестное t без такого распределения. Таким образом, $c_2 = 1597/1911 = 0.8357$. Принимая значение w_5 из § 144 за k, т. е. полагая, что $k = 0.9406$ и превосходит, как это и должно было быть (§ 118), долю c_2 осуждений, мы находим, что $P_2 = 0.9916$. Вероятность невиновности подсудимого, осуждённого и присяжными, наименьшим большинством в 7 голосов против пяти, и судьями, по меньшей мере тремя голосами против двух, стало быть, очень немного отличалась от 1/100, и из 1597 осуждённых весьма вероятно, что не подлежали осуждению примерно 15 человек.

Те же самые значения k и c_2 приводят к

$$\frac{k - c_2}{2k - 1} = 0.1188,$$

и уравнение, из которого следует определить неизвестное t, принимает вид

$$1 + 5t + 10t^2 = 0.1188(1 + t)^5,$$

так что $t = 2.789$, $u = 0.7361$. Это доказывает, что шанс u судьям не ошибаться мало отличается от найденного в § 140 для присяжных значения 0.7494 без различия категорий преступлений[13].

146. Формулы, которые мы приложили к различным проблемам суждений в уголовных делах, равным образом относятся ко всем другим видам очень многочисленных суждений, как, например, производимых полицейскими и военными судами. Но для рассмотрения таких суждений требуется, чтобы наблюдения представляли необходимые данные для установления элементов, входящих в эти формулы.

В *Comptes généraux de l'administration de la justice criminelle* содержатся и результаты, относящиеся к полицейским судам. За девятилетний период 1825 – 1833 гг. во всей Франции перед ними предстало 1 710 174 человек. Из них было осуждено 1 464 500 человек, т. е. 0.8563 из них. Это отношение мало изменялось из года в год и неизменно оставалось в границах 0.84 – 0.87. Число судей в трибуналах полицейских судов не оставалось постоянным, но их должно было быть не менее трёх, притом чаще всего их и было трое. Таким образом, двух голосов против одного было достаточно для осуждения. Полагая $n = 3$ и $i = 1$ в уравнении (118.1а), мы получим, при подстановке $t/(1 + t)$ вместо u,

$$c_1 = \frac{k(t^3 + 3t^2) + (1-k)(3t+1)}{(1+t)^3}.$$

Можно принять полученное из наблюдений приближённое и весьма вероятное значение $c_1 = 0.8563$, но этого недостаточно для определения двух неизвестных, k и t. Необходимо, кроме того, знать, сколько осуждёний из этих 1 464 500 было объявлено единогласно, и сколько простым большинством в 2 голоса против одного. Этого, однако, *Comptes généraux* не сообщают. Если предположить, что шанс не ошибаться у судей полицейских судов, как и вообще для судей, равен 3/4, и принять в предыдущем уравнении $c_1 = 0.8563$ и $t = 3$, то k окажется большим, чем единица, а принятая гипотеза недопустимой. Можно полагать, что этот шанс у судей выше, чем у присяжных, однако без необходимых наблюдений нельзя сказать, насколько.

Военный суд состоит из семи судей, и для осуждения закон требует большинства, не меньшего, чем в 5 голосов против двух. Вероятность c_2 осуждения подсудимого выводится из уравнения (118.1а) при $n = 7$ и $i = 2$ и замене u на $t/(1 + t)$. В результате оказывается, что

$$c_2 = \frac{k(t^7 + 7t^6 + 21t^5) + (1-k)(1+7t+21t^2)}{(1+t)^7}.$$

В *Comptes généraux de l'administration de la justice militaire*, публикуемых Военным министром, доля осуждённых равна по оценке 2/3. Она выведена по большому числу судебных решений, и её можно поэтому принять как приближённое и весьма вероятное значение c_2. Но для определения двух неизвестных, входящих в предыдущее уравнение, этого недостаточно. Предположив, что шанс не ошибаться у военных судей очень мало отличается от этого шанса у судей ассизов, принимая его поэтому равным 3/4, и полагая $t = 3$ и $c_2 = 2/3$, из этого уравнения следует, что $k = 0.8793$, $(1 - k) = 0.1207$. Можно ставить немного больше семи против одного за то, что военнослужащий, представший перед военным судом, виновен. Ввиду формулы (120.1) и уравнения (118.1a) мы имеем соотношение

$$(1 + t)^7 c_2 P_2 = k(t^7 + 7t^6 + 21t^5)$$

для определения вероятности P_2 вины подсудимого после его осуждения. При предыдущих значениях c_2, t, k оказывается, что $P_2 = 0.9976$. Это доказывает, что указанная вероятность мало отличается от достоверности. Полученный результат основан на условном значении t или u, степени точности которых мы не знаем, но интересно было бы быть в состоянии достоверно сравнивать вероятность суждений военной юстиции и ассизов. Для этой цели, помимо доли осуждений военнослужащих, равной 2/3, требуется знать ту же долю для единогласных осуждений и осуждений шестью голосами против одного или пятью против двух. К сожалению, этих данных нет в наблюдениях, и мы не можем предложить никакого сколько-нибудь вероятного предположения.

147. Чтобы закончить нашу книгу, нам осталось рассмотреть вероятность суждений трибуналами по гражданским делам. В гражданском процессе требуется решить, кто из тяжущихся имеет право на своей стороне. Это будет достоверно решено судьями, которые вовсе не ошибаются, и при любом их числе суждение всегда окажется единогласным. Но так совсем не бывает. Часто два равным образом просвещённых судей, самым внимательным образом исследовавшие тот же процесс, приходят к противоположным решениям. Приходится признать, что у каждого судьи существует шанс ошибаться при голосовании, т. е. не судить идеально, без всякой ошибки.

Этот шанс зависит от степени образованности и честности судьи; он заранее не известен, и, если возможно, его значение должно быть выведено из наблюдений при помощи способов, которые мы укажем. Если этот или противоположный шанс установлен для всех судей некоторого трибунала, можно будет вывести вероятность верности их суждения, или, иначе говоря, его соответствия с суждением, которое вынесли бы непогрешимые судьи. Можно также установить вероятность того, что другие судьи, также с известным шансом не ошибаться, подтвердят суждение первых.

Эта вторая задача аналогична той, которую мы изучали в уголовном судопроизводстве: величина, которую мы раньше обозначали через k, заменяется сейчас вероятностью того, что право находится на стороне одной из двух партий, определяемой первым суждением в её пользу. Но когда процесс рассматривается гражданским трибуналом впервые, не существует никакой предварительной вероятности того, что судебное решение окажется благоприятным той или другой стороне. Нет смысла определять вероятность, аналогичную k, и единственными неизвестными, которые должны быть установлены по наблюдениям, являются вероятности судей не ошибаться.

148. Рассмотрим прежде всего трибунал первой инстанции из трёх судей, A, A′ и A″. Обозначим через u, u' и u'' их вероятности не ошибаться и через c – вероятность их единогласного суждения. Последнее будет происходить, либо если никто из судей не ошибается, либо если ошибаются они все. Вероятности этих случаев равны $uu'u''$ и $(1 - u)(1 - u')(1 - u'')$, так что полное значение c равно

$$c = uu'u'' + (1 - u)(1 - u')(1 - u'').$$

После вынесения единогласного решения можно сформулировать две гипотезы и предположить [как выше], что оно было либо верным, либо нет. Первая из них означает, что никто из трёх судей не ошибся, вторая – что все они ошиблись. Вероятность наблюдённого события, т. е. единогласного решения, окажется равной $uu'u''$, если первая гипотеза верна, или $(1 - u)(1 - u')(1 - u'')$, если она ошибочна. Применяя к этим гипотезам правило о вероятности причин (§ 28), и обозначая через p вероятность первой причины, или верности решения, мы получим

$$p = \frac{uu'u''}{uu'u'' + (1-u)(1-u')(1-u'')},$$

или, иначе, $cp = uu'u''$.

Если решение не было единогласным, то один из судей проголосовал за одну сторону, а двое других – за другую. Обозначим через a, a' и a'' вероятности подобного решения, соответствующего случаям, при которых одним из судей, проголосовавшим против двух других, был A, A′ или A″. Тогда

$$a = u'u''(1 - u) + u(1 - u')(1 - u''),$$
$$a' = uu''(1 - u') + u'(1 - u)(1 - u''),$$
$$a'' = uu'(1 - u'') + u''(1 - u)(1 - u').$$

Первое, например, уравнение соответствует случаю, при котором A′ и A″ не ошиблись, а A ошибся или наоборот и аналогично остальные уравнения. Обозначим теперь через b вероятность какого-либо варианта не единогласного решения. Тогда

$$b = a + a' + a'',$$

а так как решение может быть либо единогласным, либо нет, то $b + c = 1$, что легко проверить. В результате окажется, что просто

$$b = 1 - uu'u'' - (1 - u)(1 - u')(1 - u'').$$

Для верности решения требуется, чтобы двое судей, составляющих большинство, голосуя одинаковым образом, не ошиблись, а для ошибочности решения необходимо, чтобы они ошиблись. Если поэтому обозначить через q вероятность верного не единогласного решения, то в соответствии с правилом вероятности причин или гипотез

$$bq = (1 - u)u'u'' + (1 - u')uu'' + (1 - u'')uu'.$$

Далее, пусть при очень большом числе μ суждений, объявленных одними и теми же судьями A, A′, A″, γ и β будут количествами решений, принятых единогласно и не единогласно, а среди последних пусть α, α′, α″ будут числами суждений, при которых судьи A, A′, или A″ голосовали не так, как другие двое. Тогда с очень хорошим приближением весьма вероятно окажется, что

$$\gamma/\mu = c,\ \beta/\mu = b,\ \alpha/\mu = a,\ \alpha'/\mu' = a',\ \alpha''/\mu'' = a''.$$

Число β является суммой α, α', и α'', а число b – суммой a, a' и a'' [см. выше]. Таким образом, второе из этих уравнений образовано суммой последних трёх, и потому эти 5 уравнений сводятся к четырём. Если числа α, α', и α'' известны из наблюдений, и если подставить предыдущие выражения a, a', a'' в три последних уравнения, то можно будет вывести значения u, u', u'', а при подстановке выражения для c в первое уравнение определяется γ. Если же γ также известно из наблюдений, то сравнение двух значений этой величины послужит для подтверждения теории. Поскольку u, u', u'' также определены, вероятности p и q верности единогласных и не единогласных суждений без труда устанавливаются при помощи предыдущих уравнений.

Наблюдения не указывают ни для какого трибунала чисел γ, α, α', и α'', и всё же, чтобы показать, как применять эти формулы, я произвольно выбрал вероятности $u = 4/5$, $u' = 3/5$ и $u'' = 3/5$ и предположил, что у каждого из троих судей шанс не ошибаться превышает шанс ошибки, а судьи A$'$ и A$''$ одинаково образованы, и их шансы не ошибаться одинаковы, тогда как A образован лучше, и его шанс ошибаться меньше. Тогда $c = 8/25$ и $b = 17/25$ и также $p = 9/10$ и $q = 57/85$. Можно ставить 17 против восьми или немного больше двух против одного за то, что решение судей не будет единогласным и 9 против одного и только 57 против 28 – что единогласное и не единогласное решения будут верными.

Для этих трёх судей средний шанс не ошибаться равен $(u + u' + u'')/3 = 2/3$. Если же считать их в равной мере образованными и принять эту дробь, 2/3, за общее значение u, u', u'', то окажется, что

$$c = 1/3,\ b = 2/3,\ p = 8/9,\ q = 2/3.$$

Эти значения p и q немного меньше предыдущих, и поэтому в нашем примере равная образованность трёх судей понизила вероятность верного решения и при единогласии, и без него. Но, с другой стороны, последнее значение c оказалось больше первого, а первое значение b превысило последнее, так что при этом новом условии вероятность единогласия повысилась, а потому вероятность решения большинством понизилась.

Если же мы не знаем, было ли решение троих судей единогласным или нет, основание полагать его верным будет отличаться от p и q. Обозначим в этом случае вероятность верного решения через r, тогда

$$r = uu'u'' + (1 - u)\, u'u'' + (1 - u')\, uu'' + (1 - u'')\, uu',$$

потому что при гипотезе верности решения оно, или наблюдённое событие, может произойти четырьмя различными способами, вероятностями которых являются 4 члена этой формулы. При противоположной гипотезе вероятность этого события будет равна

$$(1 - u)(1 - u')(1 - u'') + u(1 - u')(1 - u'') +$$
$$u'(1 - u)(1 - u'') + u''(1 - u)(1 - u').$$

Сумма вероятностей рассматриваемого события составляет достоверность или единицу, а делитель r, который появляется по правилу § 28, также равен единице. Заметим, что, как можно было показать непосредственно,

$$r = cp + bq.$$

Принимая предыдущие значения u, u', u'', мы находим, что $r = 93/125$. Среднее из этого значения и 2/3 оказывается равным $r = 20/27$ [19/27]. Это второе значение r немного меньше предыдущего, и поэтому верность решения, как и в предыдущем случае, стала менее вероятной при равной образованности всех троих судей.

149. Можно без труда обобщить эти формулы на случай трибунала, состоящего из произвольного числа судей, однако никакого применения нельзя будет при этом сделать ввиду недостаточности данных наблюдения, необходимых для определения шанса не ошибаться у различных судей. Если предположить, что эти шансы одни и те же, а число судей осталось равным трём, то в предыдущих обозначениях

$$c = u^3 + (1 - u)^3,\ b = 1 - u^3 - (1 - u)^3,\ cp = u^3,$$
$$bq = 3(1 - u)u^2,\ r = u^3 + 3(1 - u)u^2.$$

Принимая, далее, для c или b приближённые и весьма вероятные значения γ/μ или β/μ, можно будет из любого из двух первых уравнений определить u. Для этого достаточно знать числа γ и β единогласных и не единогласных решений среди их общего очень большого числа μ, но этих данных у нас нет. Если, к примеру, предположить, что $\gamma = \beta$, то

$$u^3 + (1 - u)^3 = 1 - 3u + 3u^2 = 1/2$$

и поэтому

$$u = [1 \pm \sqrt{3}/3]/2.$$

Одно из двух значений u превосходит 1/2, а второе меньше этого числа. Поскольку следует считать, что шанс судьи не ошибаться превышает противоположный шанс, то, принимая первое из указанных двух значений, $u = 0.7888$, мы получаем

$$p = 0.9815,\ q = 0.7885,\ r = 0.8850.$$

Пусть суждение трёх судей, единогласное или нет, пересматривается апелляционным трибуналом, скажем, из семи других судей, шанс безошибочности каждого из которых равен v. Обозначим через C вероятность подтверждения суждения первого трибунала большинством, не меньшим, чем в 4 голоса против трёх. Значение C будет дано формулой (118.1а) при замене k и u на r и v и при $n = 7$ и $i = 3$:

$$C = r[v^7 + 7v^6(1 - v) + 21v^5(1 - v)^2 + 35v^4(1 - v)^3] +$$
$$(1 - r)[(1 - v)^7 + 7v(1 - v)^6 + 21v^2(1 - v)^5 + 35v^3(1 - v)^4].$$

И если первое суждение было верным, поскольку подтвердилось вторым трибуналом, окажется, что из семи судей апелляционного суда не ошибся ни один, или ошибся 1, 2 или 3. Вероятности этих четырёх случаев соответствуют четырём членам первой квадратной скобки и поэтому их сумма, умноженная на r, равна вероятности верности и подтверждения суждения. Часть этого выражения C, умноженная на $(1 - r)$, выражает вероятность того, что суждение было ошибочно, но оказалось подтвержденным, а сумма обеих частей является полным выражением величины C. Видно также, что, обозначая через C' вероятность того, что второй трибунал отменил суждение первого, мы получим

$$C' = (1 - r)[v^7 + 7v^6(1 - v) + 21v^5(1 - v)^2 + 35v^4(1 - v)^3] +$$
$$r[(1 - v)^7 + 7v(1 - v)^6 + 21v^2(1 - v)^5 + 35v^3(1 - v)^4].$$

Суждение первого трибунала должно было быть подтверждено или отменено, и поэтому $C + C' = 1$. Это можно проверить, заметив, что сумма квадратных скобок в последней формуле равна $[v + (1 - v)]^7 = 1$. Далее, при $r = 1/2$ и произвольном значении v, или при $v = 1/2$ и произвольном значении r окажется, что $C = C' = 1/2$. Эти результаты очевидны сами по себе.

При отдельном рассмотрении двух частей выражений для C и C' можно также сказать, что первая часть C есть вероятность того, что оба трибунала решили верно, вторая часть – вероятность того, что оба они ошиблись. Те же части C' являются вероятностями того, что первый трибунал решил неверно, а второй – верно, и противоположного сложного события. И если обозначить через ρ вероятность верного решения апелляционного трибунала при верном или неверном решении первого, то эта величина окажется суммой первых частей C и C', а $(1 - \rho)$ – суммой вторых частей, что можно было установить непосредственно.

Обозначим через Γ вероятность того, что решение этого [?] суда подтверждено вторым королевским судом, состоящим также из семи судей, через Γ' – вероятность противоположного события и через w шанс не ошибаться у каждого из этих семи судей. Вероятности Γ и Γ' вычисляются по C и C' при замене r и v на ρ и w, так что

$$\Gamma = \rho^2 + (1 - \rho)^2,\ \Gamma' = 2\rho(1 - \rho).$$

Эти значения удовлетворяют условию $\Gamma + \Gamma' = 1$. Выражения для ρ и ρ' можно записать и в виде

$$\rho = \frac{r - C'}{2r - 1},\ 1 - \rho = \frac{r - C}{2r - 1}.$$

Обозначим кроме того через P вероятность верности суждения первого апелляционного суда, если оно подтверждает суждение суда первой инстанции, и через P' ту же вероятность, если первое решение отвергается. В первом случае, последовательно предполагая, что решение суда первой инстанции было верным и ошибочным, вероятность наблюдённого события, т. е. совпадения обоих суждений, будет соответственно равна первой и второй части величины C, а вероятность P первого предположения окажется равной этой первой части, делённой на сумму обеих частей. Поэтому

$$CP = r[v^7 + 7v^6(1-v) + 21v^5(1-v)^2 + 35v^4(1-v)^3].$$

Кроме того, $C'P'$ равно всей первой части C', что можно было также вывести, как это и должно было быть, из формул (120.1) и (120.2), если положить $k = r$, $n = 7$, $i = 3$. Эти уравнения можно заменить следующими:

$$CP = r\rho,\ C'P' = (1-r)\rho.$$

150. В суде первой инстанции суждение должно выноситься не менее, чем тремя судьями, а в апелляционном суде – не менее, чем семью. Эти наименьшие числа обычно не превышаются, и потому-то я и принял их. Если в моих формулах принять вместо r его значение в функции u, они станут включать два шанса, u и v, которые могут быть определены только по наблюдениям. К сожалению, известным оказывается только одна величина, доля суждений в суде первой инстанции, подтвержденных королевскими судами. Чтобы воспользоваться этими формулами, необходимо поэтому при помощи частной гипотезы свести обе неизвестные величины, u и v, к одной-единственной. Наиболее естественным представляется мне предположение, что $v = u$, т. е. равенство шансов не ошибаться у судей обоих трибуналов.

Пусть тогда из очень большого числа μ суждений суда первой инстанции было подтверждено m и, следовательно, $\mu - m$ не было подтверждено. Отношение m/μ можно принять за приближённое и весьма вероятное значение вероятности, которую мы обозначили через C, так что

$$C = \frac{m}{\mu},\ v = u,\ u = \frac{t}{1+t},\ 1-u = \frac{1}{1+t}.$$

Поэтому

$$\frac{m}{\mu} = r - \frac{(2r-1)(1+7t+21t^2+35t^3)}{(1+t)^7},$$
$$r = 1 - \frac{1+3t}{(1+t)^3},\ 2r-1 = 1 - \frac{2(1+3t)}{(1+t)^3}.$$

Подставляя эти значения в отношение m/μ, мы получим уравнение 10-й степени для определения t, а затем и u. При $v = u$ выражение для C не изменяется, если заменить u и r на $(1 - u)$ и $(1 - r)$, что соответствует замене t на $1/t$. Это означает, что если t, меньшее единицы, удовлетворяет данному значению m/μ, то этому значению равным образом будет удовлетворять и $t > 1$. Уравнение с неизвестным t относится к *возвратным* и не изменяется при замене t на $1/t$.

Принять следует значение $t > 1$, потому что оно соответствует значению $u > 1/2$, т. е. тому, что шанс не ошибаться превышает противоположный шанс, что должно иметь место при честных и образованных судьях.

151. За последние три месяца 1831 г. и за полные 1832 и 1833 гг. в публикуемых правительством *Compte généraux de l'administration de la justice civile* для каждого королевского суда указывается количества подтверждённых m и не подтверждённых $\mu - m$ суждений. Однако, лишь в юрисдикции парижского королевского суда число μ достаточно велико само по себе для того, чтобы определять t. Поэтому нам в настоящее время придётся предположить, как мы и поступили по отношению к присяжным, что шанс u не ошибаться почти одинаков для всех судей королевства. Это позволит определять t по значениям m и $\mu - m$ совместно для всех королевских судов. За последние три месяца 1831 г. и 1832 – 1833 гг. для Франции в целом

m = 976, 5301 и 5470; $\mu - m$ = 388, 2405 и 2617;
m/μ = 0.7155, 0.6879, и 0.6764

Два последних отношения, соответствующие полным годам, не отличаются друг от друга более, чем на 1/70 их среднего, что представляет собой весьма примечательный пример действия закона больших чисел[14]. Принимая за m и μ суммы чисел, относящихся ко всем трём периодам, мы получим

m = 11 747, μ = 17 157, m/μ = 0.6847.

Отдельно для парижского королевского суда

m = 2510, μ = 3297, m/μ = 0.7613.

Полученное отношение m/μ почти на 1/9 превышает его среднее значение для всей Франции.

Принимая для Франции значение 0.6847, мы находим, что

$$t = 2.157,\ u = 0.6832,\ r = 0.7626.$$

В соответствии с полученным r, не зная ни трибунала, ни сути рассматриваемого процесса, можно ставить немногим более трёх против одного за то, что суждение суда первой инстанции было верным. Заметно также, что шанс u не ошибаться у судей гражданских дел очень немного превышает дробь 0.6788 этого шанса для присяжных до 1832 г., т. е. до того, как закон предписал учитывать *смягчающие обстоятельства*.

Если принять это значение r и отношения m/μ и $(\mu - m)/\mu$ за значения C и C', из формул § 150 будет следовать, что

$$P = 0.9479,\ P' = 0.6409,\ \Gamma = 0.7466,$$

что доказывает, что можно ставить почти 19 против одного за то, что приговор апелляционного суда, подтверждающий суждение суда первой инстанции, был верным, и менее двух против одного в случае противоположного приговора.

Заметно также, что без учёта сути приговора вероятность Γ того, что он будет подтверждён вторым королевским судом на основе тех же данных, немного ниже 3/4. Значения четырех частей, составляющих выражения C и C', таковы:

$$r\rho = 0.6495,\ (1 - r)\rho = 0.2022,$$
$$r(1 - \rho) = 0.1131,\ (1 - r)(1 - \rho)=0.0352.$$

Эти дроби, составляющие в сумме единицу, выражают вероятности того, что суд первой инстанции, а затем апелляционный суд решили верно; что первый решил ошибочно, а второй верно; что первый решил верно, а второй нет; что оба решили неверно.

Примечания

1. Эту тему рассмотрел Остроградский, член Петербургской академии наук, в мемуаре, зачитанном там в июне 1834 г. Но, судя по опубликованной выдержке, которую прислал мне её автор, он исследовал эту тему совсем не так, как я. Автор

2. Умозрительные соображения бездоказательны. Шанс присяжного ошибиться несомненно меняется во времени, да и зависит от его физического и морального состояния. Оставим на совести утверждение адвоката из диккенсовых *Посмертных записок Пиквикского клуба* (гл. 34) о том, что в гражданских процессах голодные присяжные всегда решают дело в пользу истца.

3. Это трудно понять.

4. По поводу ассизов см. Прим. 7 к Предисловию.

5. Видимо, суммы (121.1) и (121.2)..

6. Одной равномерности было бы достаточно для существования линейного соотношения между X и x.

7. Нетрудное обобщение привело к сложным формулам!

8. Степень приближения не была у автора постоянной. См. также § 121.

9. В конце главы автор использовал данные и из других источников.

10. В дальнейшем автор изучал данные судов различных инстанций, и подчас трудно понять, каких именно судов.

11. Утверждение автора спорно. Относительное число преступлений в некотором регионе могло быть связано с его экономическим состоянием.

12. К статистике судопроизводства в Бельгии неоднократно обращался Кетле (Sheynin 1986, особо с. 302 – 303). Математический уровень его рассуждений был невысоким, но он, к примеру, в 1846 г. чётко указал, что после введения в стране института присяжных доля оправданий удвоилась. Он же, правда, в 1833 г. и позже утверждал противоположное; возможно, что круг лиц, из которых отбирались присяжные, был вначале очень узким.

В 1832 г. и позже Кетле приводил таблицу долей осуждения подсудимых в зависимости от их личности, в том числе от их пола.

Пуассон, в том же § 142, упомянул присяжных в Бельгии, но лишь мимоходом, и местами трудно было понять, о Бельгии или о Франции он рассуждал.

13. Приняв $t = 2.789$, мы выяснили, что левая часть этого уравнения стала равной 91.73, правая же часть – 92.78.

14. Этот закон больших чисел был вновь подтверждён значением отношения m/μ в 1824 г. Автор

Библиография

Laplace P. S. (1816), Supplément premier to *Théor. Anal. Prob. Oeuvr. Compl.*, t. 7, No. 2. Paris, 1886, pp. 281 – 320.

Sheynin O. B. (1986), A. Quetelet as a statistician. *Arch. Hist. Ex. Sci.*, vol. 36, pp. 281 – 325.