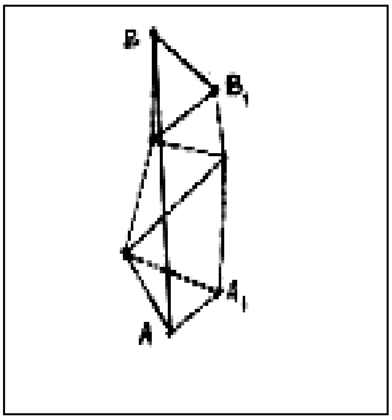
**История теории ошибок**

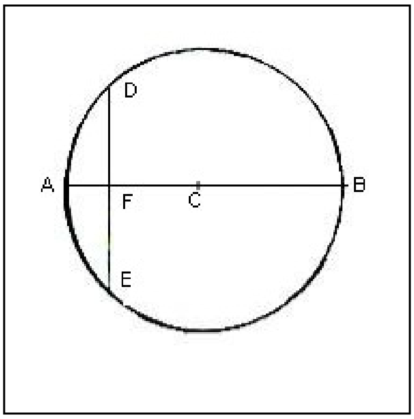
**Istoria Teorii Oshibok**

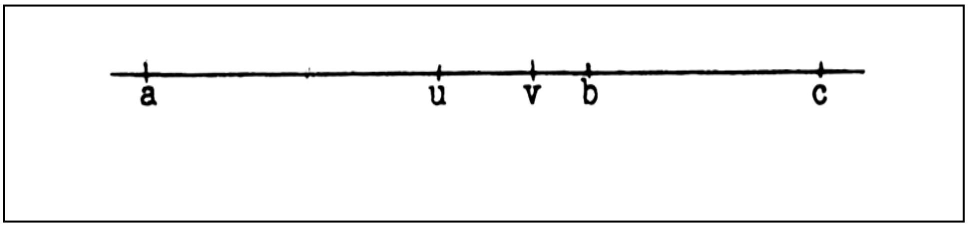
**Берлин, Berlin 2007**

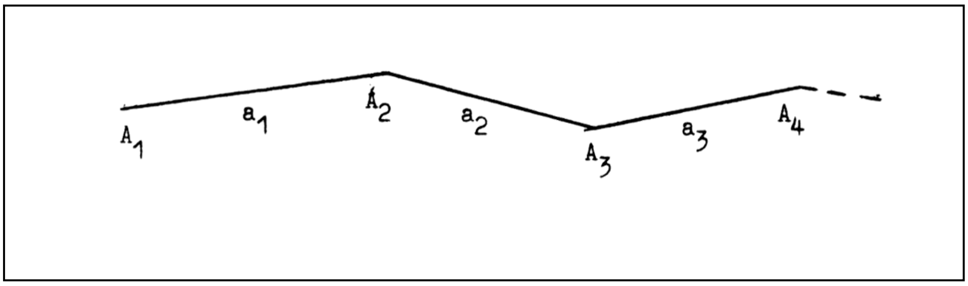
## Оглавление

**Рисунки к тексту:**



# 0. Введение

**0.1. Цели теории ошибок.** Теория ошибок это научная дисциплина, которая имеет целью определение наиболее надежных результатов измерений в экспериментальных науках. Ее можно считать соответствующим приложением статистического метода. Так, пусть даны геодезические пункты *A* и *B* и положение точки C устанавливается *засечкой,* — измерением углов *CAB* и *CBA*. Возникает несколько вопросов.

1. Какое влияние на точность определения C оказывает форма треугольника *ABC?*

2. Как и в какое время дня следует измерять углы, чтобы влияние местных метеорологических условий наименьшим образом исказило их?

3. Какое среднее значение из нескольких измерений (каждого) угла следует принять?

4. Насколько точны эти окончательные средние значения, равно как и вычисленные элементы треугольника и координаты C?

5. Если с целью контроля и повышения точности *C* засекается с трех пунктов *A, B* и *D,* так что прямые *AC, BC* и *DC,* вообще говоря, не пересекаются в одной и той же точке, то где окажется наиболее надежное положение C и насколько оно надежно?

Первый вопрос несложен. Так, ошибки неизвестных сторон треугольника могут быть вычислены по теореме синусов, а их ошибки – по дифференциальным формулам для любых заданных приближенных значений углов и их ошибок. Ответ на второй вопрос дается подходящей программой наблюдений. Например, возможно, что равное число утренних и вечерних наблюдений в наибольшей степени исключит систематические влияния изменения внешних условий.

Все остальные вопросы могут быть разрешены только стохастически и таким образом оказывается, что теория ошибок имеет две ветви, *детерминированную,* которая исследует весь процесс измерений в целом, и *вероятностную,* которая ведает обработкой их результатов.

Продолжаем рассматривать последние три вопроса. В случае одного неизвестного (*непосредственные* наблюдения) требуется выбрать его окончательное значение по наблюдениям и оценить его надежность. В общем случае, как, например, в последнем вопросе, необходимо уравнивать *косвенные* наблюдения, т. е., во-первых, вывести окончательные значения неизвестных $x, y, z, \ldots$ по избыточным системам уравнений

$$a_i x + b_i y + c_i z + \ldots + l_i = 0, \quad i = 1, 2, \ldots, n \qquad (1)$$

с заданными коэффициентами и измеренными свободными членами, и, во-вторых, оценить надежность этих значений и/или их функций. Линейность систем (1) объясняется тем, что приближенные значения $x, y, z, \ldots$ всегда известны, определяются лишь поправки к ним.

Для "физически" независимых свободных членов уравнения (1) несовместны и за решение системы приходится принимать любое множество ($\hat{x}, \hat{y}, \hat{z}$, ...), приводящее к достаточно малым остаточным свободным членам ($v_i$). МНКв не является исключением: он требует, чтобы

$$\sum v_i^2 \equiv [vv] = \min \qquad (2)$$

среди всех возможных множеств ($\hat{x}, \hat{y}, \hat{z}$, ...). И вообще выбор любого решения, который должен быть так или иначе стохастически оправдан (аналогично и в случае непосредственных наблюдений), означает наложение какого-нибудь условия на $v_i$.

**0.2. Взаимосвязь со статистикой и теорией вероятностей.** Детерминированная ветвь теории ошибок соотносится с *предварительным исследованием данных,* которое имеет целью выявление в них скрытых структур (например, систематических ошибок) и с *планированием эксперимента.* Обе названные дисциплины относятся к теоретической (но вряд ли к математической) статистике.

Далее. Случайные ошибки представляют собой частный случай случайных величин (§ 2.6), так что вероятностная теория ошибок существенно зависит от теории вероятностей, чье развитие с середины XVIII в. и, пожалуй, до 1920х годов в основном определялось необходимостью обосновывать и совершенствовать математическую обработку наблюдений. Вот, действительно, высказывание Пуанкаре (посм. публ. 1921/1980, с. 343) о своем трактате (1912, первое издание 1896): "Теория ошибок, естественно, была моей основной целью". И даже позднее Леви (1925, с. vii) заметил, что без указанной теории это (его основное сочинение, посвященное устойчивым распределениям), "не имело бы смысла".

Мы полагаем, что теория ошибок есть приложение статистического метода к обработке наблюдений в экспериментальных науках. Добавим, что теория вероятностей обобщила понятие случайной ошибки наблюдения, приняв взамен случайную величину, а математическая статистика переняла у теории ошибок дисперсию и принципы наибольшего правдоподобия и наименьшей дисперсии.

**0.3. Астрономия и геодезия.** Целесообразная обработка астрономических наблюдений позволила Кеплеру установить истинную систему мира (§ 1.9). Другой пример представляют градусные измерения, после обработки которых удалось определить общую фигуру Земли. Если некоторые точки *A* и *B* расположены на одном и том же меридиане, а *O* – центр Земли, то угол *AOB* равен разности измеряемых широт этих точек, а длина дуги *AB* косвенно измеряется при помощи цепи треугольников, т. е. по измерению двух (практически – всех трех) углов каждого ее треугольника и длины некоторого базиса (фактически – двух базисов, в начале и конце цепи). Градусное измерение и состоит в совокупности обоих указанных измерений. **рисунок**

В древности ученые пытались определить радиус сферической Земли, но после того, как Ньютон доказал, что Земля представляет собой эллипсоид вращения, стало необходимым определять его параметры, – большую полуось *a* и сжатие $(a - b)/a$. Для этого достаточно двух градусных измерений, но для более или менее точного установления параметров эллипсоида требуется много больше. В конце XVIII в. градусные измерения потребовались и для перехода к метрической системе мер (1 метр = $10^{-7}$ четверти меридиана), и для нужд картографирования (которое через несколько десятилетий начало основываться на триангуляции вообще). Маятниковые наблюдения явились важным дополнительным средством для установления формы (но не размеров) Земли, а точнее, ее сжатия. В наше время они, наряду с другими средствами, служат для изучения гравитационного поля Земли.

Заметим, однако, что приложения теории ошибок вовсе не исчерпываются областями указанных наук (§ 6.6.1).

**0.4. Когда и почему возникла теория ошибок.** Кноблох (1990, с. 307) указал, что теория ошибок возникла ввиду стремления объяснять природу в математических терминах и необходимости повысить точность экспериментальных наук, равно как и вследствие развития теории вероятностей. Мы возражаем против последнего; именно теория ошибок побуждала совершенствовать теорию вероятностей (§ 0.2).

Перечислим также более непосредственные условия, оказавшиеся необходимыми для возникновения теории ошибок. Нужно было изучить суть и характер влияния случайных и систематических ошибок на наблюдения, сформулировать основные цели и методы обработки наблюдений и описать и то, и другое в отдельном сочинении. Соответственно, мы полагаем, что теория ошибок возникла в XVIII в. и попробуем теперь выделить ее стадии развития, отдельно для обеих ее ветвей, начиная с вероятностной.

*Первый период.* Ученые относились к своим наблюдениям как к частной собственности, так что большая часть данных видимо оставалась не известной научному миру. Птолемей может служить здесь примером, тогда как Тихо Браге возвестил начало нового периода.

*Второй период.* Все наблюдения по крайней мере должны были сообщаться, однако их обработка оставалась либо субъективной, либо, в лучшем случае, без достаточного вероятностного обоснования (Бошкович).

*Третий период.* Обработка наблюдений начала сопровождаться утверждениями о вероятностных или статистических свойствах окончательных результатов, а позднее эти свойства стали широко известны. Этот период начался с Бошковича или немного раньше. Во второй половине XVIII в. были введены первые вероятностные распределения, частично обосновано применение среднего арифметического, предложен принцип наибольшего правдоподобия, началось оценивание точности, возникла и теория ошибок, и сам этот термин, началось уравнивание косвенных наблюдений и предвосхищен принцип наименьших квадратов.

*Четвертый период.* Была создана классическая теория ошибок (Лаплас, Гаусс), выполнены существенные дополнительные исследования (Гельмерт), теория ошибок была методологически улучшена, делались попытки обобщить и усовершенствовать результаты Гаусса.

Сколь ни важен был вклад Лапласа, именно Гаусс создал МНКв как практический аппарат. В отличие от первого, он не предполагал большого числа наблюдений и действительно дал возможность наблюдателям практически использовать свои результаты. Вот мнение Субботина (1956, с. 297), об определении орбит небесных тел, полностью применимое и к нашей теме: Лагранж и Лаплас

> *Ограничились лишь математической стороной дела, тогда как Гаусс не только тщательно обработал свое решение с точки зрения вычислительной техники, но и учел все условия работы и все привычки астрономов-вычислителей.*

В истории детерминированной ветви теории ошибок мы выделяем

*Период до XVIII в.* Стали известны существенные источники ошибок наблюдений и общие меры, предохраняющие от их влияния или сводящие его к минимуму.

*XVIII в. и Лаплас.* Появились дифференциальные формулы оценки точности геодезических сетей, на этой же основе начало изучаться влияние ошибок наблюдений и инструментальных ошибок, были явно выделены понятия о систематических и случайных ошибках.

*Гаусс и Бессель.* Началась охота на ошибки. И приборы, и методы наблюдений стали считаться негодными, пока и поскольку не заканчивалось их тщательное исследование и не обнаруживались меры по устранению (по сведению к минимуму) выявленных погрешностей.

*Гельмерт.* Достигнуто более обширное исследование точности геодезических сетей. Их отдельные части временно, перед началом общего уравнивания, могли быть теперь заменяемы геодезическими линиями.

**0.5. Содержание книги.** Будучи по образованию инженером-астрономом-геодезистом и математиком, мы смогли хорошенько изучить нашу тему, которой посвятили немало статей. Одну из них (Шейнин 2000а) мы рекомендуем в качестве сводки этой книги, которая является дополненным и переработанным вариантом английского издания (1996а).

Отправляясь от многих источников, в том числе от наших собственных статей, мы описываем предысторию и саму историю теории ошибок от Птолемея до Кеплера и Галилея, периода градусных измерений, до Ламберта, Симпсона, Лапласа, Гаусса и Бесселя и, наконец, до начала XX в.

Из предшествующей литературы мы прежде всего назовем Гельмерта (1872) и Идельсона (1947). Обе эти книги, конечно же, устарели, но по-прежнему полезны, а со второй из них началось и активное вторжение в теорию ошибок математико-статистических идей. Мы также рекомендуем сочинение Фейербрадера (1999), написанное с современной точки зрения.

Здесь же упомянем в противоположном ключе книгу Стиглера (1986), который первым (и, надо думать, последним) посмел значительно принизить заслуги Гаусса и даже оклеветать величайшего ученого, а кроме того, не разобравшись, обвинил Эйлера в непонимании целей уравнивания наблюдений. В последующей работе Стиглер (1999, с. 317 – 318), нисколько не смущаясь, назвал Эйлера крупным статистиком. Стиглер (там же, с. 52) допустил несколько других грубейших искажений, а одного автора, который заявил, что Пуассон ввел усиленный закон больших чисел (см. нашу рецензию в *Isis*, vol. 92, 2001, pp. 184 – 185), он (там же) назвал первоклассным ученым. Я оказался единственным автором, который, вопреки утверждению Хальда (1998, с. XVI), назвавшего книгу Стиглера (1986) *эпохальной*, опроверг его измышления, см., например, Шейнин (1999a; 1999b).

Надеемся, что эта книга окажется полезной тем, кто интересуется историей математики или экспериментальных наук и, конечно, статистики. От читателей требуется некоторое знание математического анализа и классической теории вероятностей.

**0.6. Терминология и обозначения.** Без потери общности мы ограничиваемся в системах (1) случаем двух или трех неизвестных и называем эти системы линейными, опуская другие прилагательные (избыточная, алгебраическая). Два термина, случайная величина и нормальный закон, мы применяем вне зависимости от момента их появления (§§ 2.5 и 4.6.2). Мы также используем стандартное обозначение $\bar{x}$ для среднего (выборочного) арифметического из наблюдений $x_i$ и E$\xi$ для математического ожидания случайной величины $\xi$. Наконец, мы применяем исключительно удобное и изящное обозначение, введенное Гауссом (1811, § 13)

$$[ab] = a_1b_1 + a_2b_2 + \ldots + a_nb_n,$$

ср. обозначение $[vv]$ в условии (2). Заметим, что Лаплас не воспринял его и что, хотя этот великий ученый и не мог служить образцом методического изложения своих результатов, позднейшие французские ученые также не последовали за Гауссом.

В том же сочинении Гаусс ввел и аналогичные обозначения, которые потребуются нам в § 5.1.6:

$$[bc, 1] = [bc] - [ab][ac]/[aa],\ [cd, 2] = [cd, 1] - [bc, 1][bd, 1]/[bb, 1]$$

и т. д. Знаменатель второго члена в последнем выражении начинается с буквы, предшествующей *c,* т. е. с *b* и потому равен [*bb,* 1], поскольку 1 предшествует числу 2 в левой части равенства. Второй сомножитель в числителе того же члена содержит "произведение" *bc,* потому что *c* – первая буква в [*cd,* 1], а *b* –первая в [*bb,* 1]. Аналогично, [*bd,* 1 ] появилось потому, что *b* и *d* – вторые буквы в [*bb,* 1] и [*cd,* 1] соответственно.

Эти обозначения, притом уже без запятых, удобны при решении нормальных уравнений по когда-то универсальному, но и теперь не забытому методу Гаусса последовательного исключения неизвестных.

*Признательность.* Историю теории ошибок я начал исследовать в 1962 г., но свои первые статьи я в Библиографию не включил, поскольку больше не считаю их достойными.

Многие выдающиеся ученые так или иначе помогали мне. Покойный А. П. Юшкевич хорошо относился ко мне и представил серию моих рукописей в *Archive for History of Exact Sciences,* чьему редактору, также покойному Клиффорду Трусделлу, пришлось-таки повозиться с моим английским. Со временем я смог су-

щественно улучшить свои знания языка и начал всерьез заботиться о стиле (также и в своих русских статьях). Затем один из моих корреспондентов указал мне, что А. П. слишком доверял моим писаниям, и я перестал просить его представлять мои работы.

Да, в конце 1960-х годов я решил публиковаться в основном за рубежом, хотя бы уже потому, что в Москве мне стало тесно. Вначале это было сравнительно легко, затем порядки стали ужесточаться и две мои рукописи были без должного разрешения вывезены на Запад советским и американским коллегами соответственно, но назвать их мне всё еще не хотелось бы.

После эмиграции в 1991 г. мне удалось возобновить свои изыскания, в большой степени ввиду моральной и научной поддержки И. Пфанцагля, профессора Математического института в Кёльне, который даже смог каким-то образом выхлопотать для меня (давно иссякший) грант от издательства Акселя Шпрингер.

Я неоднократно пользовался советами покойных Ч. Эйзенхарта и У. Краскла (который также связал меня с Пфанцаглем), а также Р. У. Фейрбрадера, Р. Л. Плакетта и покойного М. В. Чирикова, талантливого математика, чья научная карьера была рано оборвана заболеванием.

## 1. Ранняя история

С современной точки зрения подход древних астрономов к обработке наблюдений трудно объясним; мы попытались систематизировать и прояснить эту тему.

**1.1. Границы и оценки.** Архимед (1925, с. 68 – 69) был, видимо, одним из первых, кто заявил, что ни человеческие способности, ни приборы не обеспечивают достаточной надежности наблюдений. Однако, продолжал он, поскольку эта тема "часто" обсуждалась, нет необходимости входить в подробности. Нам, к сожалению, ничего не известно об этих более ранних высказываниях.

Понимая, что их наблюдения несовершенны, древние ученые пытались определить границы для измеряемых величин. Эта практика "стала хорошо известным приемом, ... который применяли, например, Аристарх, Архимед и Эратосфен" (Тумер 1974, с. 139). Для примера мы процитируем Аристарха (1959, с. 403):

*Отношение диаметра Солнца к диаметру Земли больше, чем ...* [19:3 = 6.33], *но меньше, чем ...* [43:6 = 7.17].

Пусть наблюдения неизвестной константы $A$ обозначены $x_1$, $x_2$, ..., $x_n$, $x_1 \le x_2 \le \ldots \le x_n$. За искомые границы можно было бы принять $x_1$ и $x_n$, однако с возрастанием числа наблюдений $x_1$ будет вероятно убывать, а $x_n$ – возрастать, так что границы остаются неопределенными и их придется, видимо, устанавливать по косвенным данным или теоретическим соображениям. Опять же, границы вряд ли окажутся полезными, если наблюдаемая величина должна послужить исходным параметром или аргументом в трудных вычислениях. Положение осложнится ещё более, если придется прибегать к нескольким подобным величинам[1].

Другими словами, пределы не исключали необходимости в назначении какой-либо точечной оценки для $A$. Древние астрономы вряд ли применяли универсальную оценку (например, среднее арифметическое); представляется, что они каждый раз выбирали какое-то число с учетом предыдущих наблюдений, субъективных восприятий, равно как и удобства последующих вычислений (Нейгебауэр 1950, с. 252). Быть может только один раз Птолемей (I 12, с. 63; Н 68)[2] пояснил свой выбор. Он остановился на ( $x_1$ + $x_n$)/2, так что наибольшая возможная ошибка оказалась наименьшей, ср. метод минимакса в § 1.9. Было известно, как он заявил, что удвоенное наклонение эклиптики более 47°40' и менее 47°45' и что поэтому

> *мы выводим почти то же отношение, что и Эратосфен, которое принял и Гиппарх. Ибо* [в соответствии с этим] *дуга ... примерно равна* (11/83) 360° = [47° 42' 39"].

И он заметил, что "принимает точку на полпути между двумя крайними"[3].

**1.2. Регулярные наблюдения.** Другой важной особенностью древней астрономии было понимание необходимости регулярных наблюдений. Птолемей (III 1, с. 132; Н 194) засвидетельствовал, что Гиппарх регулярно наблюдал длину тропического года. Мало что известно про этого ученого, труды которого существенно помогли Птолемею разработать свою классическую (но неверную) систему мира и которого последний (IX 2, с. 421; Н 210) назвал "великим приверженцем истины". Это многозначительное замечание, видимо, означает, что Гиппарх не опасался сообщать о расхождениях между наблюдениями, ср. Тумер (1974, с. 140).

Остается неизвестным, в какой мере Гиппарх представлял себе, что регулярные наблюдения уменьшают влияние некоторых (случайных) ошибок и могут исключать систематические влияния.

**1.3. Наилучшие условия для наблюдений.** Третьей и последней особенностью древней астрономии, которую Нейгебауер (1950, с. 250) приписывает даже вавилонским астрономам эпохи Селевкидов, было использование наилучших условий наблюдения. Так, в некоторые периоды времени ошибки заданной величины в регистрации момента астрономического события влияют на окончательный результат намного меньше, чем в другое время. Или же (Птолемей V 14, с. 252; Н 417) равенство каких-либо двух углов можно было установить легче, чем их величины. Здесь наилучшие условия связывались с возможностью полного отказа от измерений, неизбежно искаженных существенными ошибками.

Нейгебауер (1948, с. 101) также заметил, что наблюдения в древности были "скорее качественными, чем количественными". Мы не удовлетворены этим выражением, который другие авторы (Аабо и Де Солла Прайс 1964) даже использовали в заглавии своей статьи, потому что почти все наблюдения являлись количественными. И всё же следует добавить, что, вообще говоря, древняя наука, в отличие от новой, была качественной и что выбор наилучших условий для наблюдений является одной из целей детерминированной теории ошибок (§ 0.1).

Действительно. Пусть результат наблюдения, *y,* зависит не только от своего аргумента, но и от нескольких параметров (условий наблюдения). Тогда его дифференциал *dy* может быть вычислен для любого набора параметров и их погрешностей и это означает, что астроном может заранее установить *dy* и выбрать разумные условия наблюдения. В древности определение точности подобных функций было возможно лишь методом проб и ошибок и, как представляется, астрономы того времени более или менее успешно пользовались им. И неудивительно, что они часто отбирали лишь одно или несколько наблюдений, используя остальные лишь для грубого контроля. Так, Птолемей (III 1, с. 137; H 203) "отставил" наблюдения, "выполненные довольно грубо", а Бируни (1967, с. 46 – 51) отбросил 4 косвенных наблюдения широты некоего населенного пункта и воспользовался только одним, простым и непосредственным.

**1.4. Птолемей.** Он в первую очередь был не наблюдателем, а теоретиком. Тем не менее, в его творчестве прослеживается по крайней мере две из трех особенностей древней астрономии. Во-первых, он наблюдал регулярно (Птолемей III 1, с. 132 и 136; H 194 и 201), Кеплер же (1609/1992, с. 647/327) "неоднократно отмечал", что Птолемей имел в своем распоряжении "намного больше наблюдений, чем представил в своем труде". И безусловно ясно, что, во-вторых, Птолемей представлял себе всё, что можно было о наилучших условиях наблюдений. Вот одно из его утверждений (IX 2, с. 423; H 213):

> *Наиболее вероятно* (most likely), *что надёжны те наблюдения планет, при которых было отмечено* [их] *действительное* [видимое] *соприкосновение или сильное сближение со звездой или Луной, и особенно те, которые выполнены при помощи астролябии.*

Вряд ли он отклонялся от своих же (косвенных) рекомендаций и обманщиком он никак не был, тем менее "самым успешным обманщиком в истории науки" (Р. Р. Ньютон 1977, с. 379), выдумавшим все свои наблюдения[4]. Да, многое можно добавить о непонятном поведении Птолемея. Кеплер (1609/1992, с. 642/324) полагал, что

> *Вряд ли мы имеем что-либо от Птолемея, в чем нельзя было бы достаточно обоснованно усомниться прежде, чем оно окажется полезным для нас в достижении требуемой точности.*

Верные Птолемею современные астрономы признают, что он отбрасывал, подравнивал или включал без изменения "громадное множество" данных "как только считал нужным" (Джинджерих 1983, с. 151), что он был оппортунистом, готовым "упрощать и стряпать"(Уилсон 1984, с. 43). И вот забытое утверждение Ньюкома (1878, с. 20): "Весь *Альмагест* [основное сочинение Птолемея], ... как мне представляется, дышит безупречной искренностью".

Особо остановимся на забытом также мнении Лапласа (1796/1884, с. 421 – 423). Альмагест – "склад древних наблюдений, один из наиболее ценных памятников древности". Птолемей "занимает достойное место в истории науки". Он очень часто ссылается на Гиппарха "в почетной манере" (см. также наш § 1.2) и уже поэтому не мог приписывать себе его каталог. Он оказал большие услуги географии, в частности сбором географических координат различных мест. И вот последнее высказывание (с. 423):

*Его репутация испытала ту же судьбу, что и у Аристотеля и Декарта; их ошибки не были известны, но затем от слепого восхищения перешли к несправедливому пренебрежению, ибо даже самые полезные перевороты в науке ни в коей мере не исключают страстей и несправедливости.*

Трудно понять подход Птолемея к округлениям[5]. Комментаторы, видимо, согласны в том, что и вообще (Нейгебауер 1950, с. 252)

*Числа несомненно улучшались для облегчения вычислений, и это видно на бесчисленных примерах в греческой и вавилонской астрономии. Часто заметно округление промежуточных результатов, равно как и важных параметров, что нередко лишает нас всякой надежды точно воспроизвести исходные данные.*

И снова он же (1975, с. 107):

*По всей древней астрономии непосредственные наблюдения и теоретические соображения безнадежно переплетены ... неизбежно имеющие место числовые неточности и произвольные округления ... то и дело имеют тот же порядок, что и исследуемые величины*[6].

Возможно, что приближенные вычисления относились к *низшей*, прикладной, а не *научной* математике (Лурье 1934, с. 37).

Птолемей безусловно имел общее представление о систематических и случайных ошибках. Иначе быть просто не могло, да и сам он (VIII 6, с. 416; H 203) указал, что различия

*Между самими наблюдателями и* [состояниями] *атмосферы в районах наблюдения могут привести к изменениям и сомнению в моментах* [явления], *как это стало ясно, по крайней мере мне, из моего собственного опыта и ввиду расхождений в этого рода наблюдениях.*

И вот другое высказывание Птолемея (1956, III, 2, с. 231):

*Практически все другие гороскопические приборы, на которые надеется большинство более осторожных наблюдателей, часто допускают ошибки; приборы для наблюдения Солнца – ввиду случайного смещения своего положения или гномона, а водяные часы – вследствие засорения или неравномерного истечения воды из-за различных причин и просто случайно.*

**1.5. Некоторое пояснение.** Постараемся сформулировать несколько принципов, которыми быть может руководствовались древние астрономы.

1. Они относились к своим наблюдениям как к частной соб-

ственности и отбрасывали некоторые из них, не сообщая об этом никому.

2. Оставшиеся наблюдения становились известными другим и могли использоваться ими даже без всяких ссылок; каждый знал, чем занимался или занимается другой. Многие авторы, начиная по крайней мере с Бируни (Шевченко 1988, с. 175), поверили в то (а некоторые продолжают верить и сейчас), что Птолемей приписал себе звездный каталог Гиппарха. Мы полагаем, что не приписал, а (возможно) переписал, притом с чистым сердцем.

3. Понимая, что, несмотря на принятые меры, (большинство) наблюдений оставались искаженными существенными ошибками, астрономы считали неиобходимым подправлять полученные результаты. Иначе говоря, они рассматривали результат наблюдения не как число или точку, а почти как любое число, расположенное внутри оцененных ими пределов (§ 1.1).

Сейчас известно, что допустимо оставлять любое из подобных наблюдений и отбросить все остальные. Вспомним, действительно, о распределении Коши, при котором среднее арифметическое не лучше отдельного наблюдения, к тому же подобная практика соответствовала бы качественному характеру древней науки.

Картографические работы Птолемея косвенно подтверждают последний вывод (Берггрен 1991): он в основном стремился к подобию правды (мы бы сказали: к истине в целом), а не к математической согласованности результатов[7]. И вот подобное же суждение (Прайс 1955, с. 6) о намного более близком периоде:

*Многие средневековые карты вполне могли быть составлены, исходя из общего знания местности, без всяких измерений.*

**1.6. Бируни.** Он был единственным арабским ученым, который превзошел Птолемея и оказался достойным предшественником Галилея и Кеплера. Естественно, что он придерживался разумных традиций древней астрономии.

**1.6.1.** Он (1967, с. 203) по крайней мере стремился установить пределы искомых величин:

*Что касается широты Багдада, различные наблюдатели установили, что она не меньше 33° 20′ и не больше 33° 30′; общепринятое значение – 33° 25′, потому что оно также является средним между ними.*

Представляется, что полусумма крайних наблюдений начала признаваться как возможная оценка, однако *также* подразумевало, что в ходу были и другие оценки, не обязательно среднее арифметическое, см. также ниже. Вот еще два случая:

*Что касается среднего из двух моментов, то оно является общим правилом, принятым многими вычислителями для возможного уменьшения ошибок наблюдения, так что вычисленный момент оказывается между верхним и нижним пределами* (с. 168).

А на с. 237 Бируни замечает, что будет "доверять" определенной величине,

*потому что она близка к средней из меньшей …и большей величин …и потому что косвенный метод... приводит к ненамного отличающейся величине.*

**1.6.2.** Бируни (1967, с. 32) рекомендует "постоянно" наблюдать широты, чтобы предупреждать об оползнях и т. д. Но при неуверенности примерно в 1′ (около 2*км),* см. § 1.6.1, подобные наблюдения, не говоря уж о долготах, вряд ли решат поставленную задачу. Впрочем, в том же источнике Бируни несколько раз упоминает регулярные наблюдения, а на с. 65 указывает, что "Ал-Баттани заявил, что повторял свои наблюдения в течение многих лет".

**1.6.3.** Приведем лишь один отрывок (с. 58) об использовании наилучших условий наблюдений: “Первый метод надежнее, потому что он зависит от непосредственных наблюдений и не нуждается ни в каких вычислениях”.

Бируни (1967, с. 152) был, видимо, первым, кто оставил рассуждение (правда, лишь качественное) об ошибках вычислений и об их совместном действии с погрешностями наблюдений:

*Употребление синусов порождает погрешности, которые становятся заметными, если они присоединяются к ошибкам, вызванным применением малых инструментов, и к погрешностям, допускаемым наблюдателями.*

Следует остановиться еще на выборе оценки для непосредственных наблюдений и на соображениях об исключении систематических влияний.

**1.6.4.** По крайней мере в одном месте Бируни (1967, с. 83) выбрал удобную и разумную, хотя вряд ли общепринятую оценку:

*Все собранные нами свидетельства совместно указывают, что* [наклонение эклиптики] *составляет 23º плюс одна треть и одна четверть градуса. Небольшой избыток или недостаток в некоторых результатах* [наблюдений] *вызваны прибором.*

Мы подтверждаем только что сделанный вывод тем, как Бируни (1983, см. Ал-Хазини 1983, с. 60 – 62) вычислял плотность металлов: иногда он принимал моду, но в других случаях – либо полусумму крайних измерений, либо какую-нибудь величину между ними.

**1.6.5.** Описывая определение разности долгот двух городов, Бируни (1967, с. 155) рекомендовал:

*Наблюдатели* [лунного] *затмения определяют все его моменты* [фазы], *так что каждый из них в одном городе может быть соотнесен с соответствующим моментом в другом. Также из каждой пары противоположных моментов следует определять средний момент затмения.*

Пусть наблюдатели в двух городах отметят 5 фаз затмения, $u_i$ и $v_i$, $i$ = 1, 2,..., 5. Тогда величины $\Delta\lambda_i = u_i - v_i$ следует сравнить друг с другом и неизвестное $\Delta\lambda$ будет как-то определено. "Средние моменты" окажутся равными $u_3$, $(u_1 + u_5)/2$ и $(u_2 + u_4)/2$ и соответственно во втором городе, причем сравнение разностей между тремя значениями этих моментов для обоих городов может обеспечить некоторые сведения о том, определяли ли наблюдатели соответствующие фазы одним и тем же образом. Если и только если это имело место, указанные разности характеризовали бы случайную составляющую погрешности.

**1.7. Галилей.** Майстров (1964; 1967, § 5 в гл. 1) первым описал мысли Галилея (1632, День третий), относящиеся к обработке наблюдений параллакса Новой звезды 1572 г. Впрочем, Буняковский (1846) посвятил этой теме несколько строк в главе об истории теории вероятностей, но не указал точной ссылки. Более подробное и точное изложение, чем у Майстрова, см. Хальд (1990, с. 149 – 160).

Способ определения параллакса был явно негодным: в те времена даже годичные (а не суточные, которые астрономы тогда пытались измерить) параллаксы не поддавались оценке. Астрономы, правда, интересовались лишь местом Новой (*под* Луной или *среди* неподвижных звезд), но это облегчение по существу ничего не меняло. Интересно, однако, высказывание Галилея (с. 214):

> *Самым подходящим будет внести поправки и исправления, наименьшие и наиболее близкие, какие только возможно ... если можно смягчить явную ошибку... прибавлением или вычитанием двух или трех минут ... то не следует стремиться исправлять их добавлением или отнятием 15, 20 или 50 минут.*

См. § 2.4.1 по поводу аналогичной рекомендации Бошковича. Галилей кроме того указал два свойства "обычных" случайных ошибок, а именно, что меньшие по абсолютной величине ошибки более вероятны и что ошибки, противоположные по знаку, равновозможны.

Галилей (1613/1957) также изучал поведение солнечных пятен и, сумев отделить их случайное движение от движения вместе с солнечным диском, тем самым определил период обращения Солнца около своей оси. Заметим кстати, что по свидетельству Марко Поло существование солнечных пятен стало известно китайским астрономам не позже конца XIII в. (Шейнин 2005b). Не зная этого, Гумбольдт ( /1858, т. 4, с. 64 прим.) тем не менее считал возможным их столь раннее обнаружение.

**1.8. Тихо Браге.** Уэсли (1978, с. 52) утверждает, что Браге был первым, кто понял, что необходимо получать длинные ряды наблюдений, чтобы случайные, инструментальные и человеческие ошибки могли уравновеситься. Вторую половину фразы следует просто отбросить, настолько она поверхностна, первая же важна. Впрочем, регулярные наблюдения проводились и в древности (§§ 1.2, 1.4 и 1.6), но Браге был их первым последователем в новое время.

Уэсли (с. 51) также замечает, что Браге сочетал наблюдения, выполненные при помощи нескольких инструментов. Мы не знаем, как он поступал, когда какой-либо из них временно не был в употреблении, т. е. когда осредненные значения измеряемых величин возможно сдвигались, и как он учитывал результаты менее надежных инструментов. Во всяком случае, непрестанное стремление Браге к достижению наивысшей возможной точности и его очевидный успех в этом направлении дали возможность Кеплеру создать верную систему мира.

Сам Браге разработал "промежуточную" и ныне забытую систему, в которой Солнце вместе с остальными планетами обращались вокруг Земли. При этом он должен был как-то уравнивать наблюдения, но как именно остается неизвестным. Впрочем, один эпизод можно описать (Плакетт 1958/1970, с. 122 – 123). Браге объединил попарно 24 своих наблюдений прямого восхождения некоторой звезды и вычислил средние арифметические каждой пары. Далее он соединил воедино полученные таким образом значения с тремя одиночными наблюдениями, придав всем 15 одинаковый вес (и тем самым косвенно воспользовался обобщенньм средним арифметическим). Пары он выбрал так, чтобы по возможности исключить из частных средних систематические погрешности и, видимо, чтобы оценить, правда, лишь качественно, влияние случайных ошибок в 12 случаях из 15, ср § 2.5.1. Возможно, что в одиночные наблюдения он предварительно ввел какие-то поправки[8].

**1.9. Кеплер.** В отличие от Птолемея и Бируни Кеплер (1606/2006, с. 163) сформулировал свое мнение о случайности:

*Но что такое случайность? Всего лишь идол, и притом самый отвратительный из идолов; ничто, кроме как оскорбление полновластного и всемогущего Бога, равно как и совершеннейшего мира, который вышел из Его рук. Вместо души случай обладает опрометчивыми побуждениями, а вместо тела – безграничным хаосом. И кощунственно приписывать ему божественную вечность и всемогущество и божественное сотворение мира.*

И всё-таки его законы движения планет не могли обосновать значений эксцентриститетов их орбит, так что в конце концов Кеплер был вынужден признать их случайными, вызванными возмущающими причинами.

При обсуждении и обработке наблюдений Кеплер не отделял систематических ошибок от случайных, но его доводы в основном относились к последним. Так, он (1609/1992, например с. 210/63, 215/71) заявил, что ошибки неизбежны. В одном случае (с. 286/113) Кеплер на равных основаниях применил термины *неуверенность* и *latitude* (очевидно *размах*) наблюдений. Размах вероятно возрастает с числом наблюдений (см. также § 1.1) и поэтому не является надежной оценкой точности, но ведь ничего лучшего в те времена не было известно. Далее, Кеплер (с. 520/254) намекнул на равную вероятность ошибок обоих знаков: "если ... мы берем среднее ... как бы говоря, что в двух наблюдениях ... были какие-то небольшие ошибки ... противоположных знаков ..."

Здесь заметен и подход к обоснованию среднего арифметического, хотя только в простейшем случае двух наблюдений. Косвенное и более общее замечание того же смысла, притом относящееся к предыстории закона больших чисел, содержится в его письме 1627 г. (Каспар и фон Дик 1930, с. 248): общий вес большого числа монет одной и той же чеканки, как он утверждал, не зависит от неточностей в весе отдельных монет. Обозначим вес монеты $i$ через $a_i$, $i = 1, 2, ..., n$. Тогда утверждение Кеплера, но только в исправленном виде, означало, что при большом $n$ $\sum a_i /n \approx$ Const.

Кеплеру пришлось проделать громадные вычисления и, в частности, уравнивать и прямые, и косвенные наблюдения. Вот самый интересный пример в первом случае (Кеплер 1609/1992, с. 200/63). Имея наблюдения

23'39", 27'37", 23'18", 29'48",

он принял за окончательное значение 24' 33" как "среднее по добру и справедливости" (*medium ex aequo et bono*). Реконструкция (Эйзенхарт 1976, с. 356) такова: это значение является обобщенным средним арифметическим с весами наблюдений соответственно 2, 1, 1 и 0 (последнее наблюдение отброшено). Но еще интереснее, что приведенная им фраза встречалась у Цицерона (*Pro A. Caecina oratio*, § 65)[9], которого Кеплер наверное читал, и имела оттенок "А не в соответствии с буквой закона", см. список юридических изречений и выражений, Розенталь и Соколов (1956, с. 126). Иначе говоря, среднее арифметическое стало *буквой закона.*

По крайней мере в одном случае метод Кеплера (1609/1992, с. 521 – 524/255 – 256) уравнивания косвенных наблюдений не был достаточно общим; при других исходных данных ему, видимо, пришлось бы поступать иначе. Вот, например, его замечание (с. 523/256):

> *Поскольку первое и третье положения Марса ... согласуются друг с другом довольно хорошо, некоторые менее мыслящие лица подумали бы, что его* [окончательное положение] *следует устанавливать по ним, а другие как-то приводить в соответствие с ними. И я сам довольно долго пытался добиться этого, но поскольку* [мне это не удалось], *их пришлось также учесть.*

Несомненно, впрочем, что Кеплер не грешил против здравого смысла и удерживал поправки "в границах точности наблюдений" (1609/1992, с. 334/143). И его знаменитое высказывание (с. 286/113), ознаменовавшее опровержение птолемеевой системы мира, подтверждает только что приведенную фразу:

> *Благость Божья соизволила дать нам в лице Тихо столь прилежного наблюдателя, наблюдения которого указывают на ошибку в 8' в этом вычислении по Птолемею ... поскольку ими нельзя пренебречь, уже одни эти восемь минут указали путь к*

*преобразованию всей астрономии и доставили материал для большей части данной работы.*

Мы полагаем, что Кеплер применил элементы метода минимакса (§ 4.4), в соответствии с которым наибольший по абсолютной величине остаточный свободный член заданной системы уравнений должен быть наименьшим из всех ее "решений". Кеплер, видимо, отыскивал этот минимум среди нескольких разумных вариантов и убедился в невозможности сочетать птолемееву модель с наблюдениями Браге[10].

Принцип минимакса имеет отношение к теории принятия статистических решений, но не к вероятностной теории ошибок. Интересно, однако, что он соответствует обобщенному МНКв (§ 4.4) и что в случае непосредственных наблюдений он приводит к выбору среднего из крайних наблюдений (а не к среднему арифметическому). Даниил Бернулли (1778, § 10)

*Выяснил, что в качестве правила для нескольких наблюдений это менее часто ошибочно* [видимо: менее уклоняется от среднего арифметического], *чем я* [он] *полагал до соответствующего исследования.*

Никаких подробностей он, однако, не привел, а его исследование (и соответствующая публикация?) остается неизвестным[11].

Возвращаясь к Кеплеру, заметим, что иногда (также, видимо, в указанном выше случае со с. 334/143) он уравнивал наблюдения, искажая их произвольными малыми поправками, которые должны были соответствовать свойствам "обычных" случайных ошибок и представляется, что он таким образом применял элементы метода Монте Карло.

**Примечания**

**1.** Подобные задачи были бы трудны и сейчас, поскольку аргументы могут оказаться зависимыми друг от друга, и/или же их вероятностное поведение существенно различным.

**2.** Мы ссылаемся на английское издание Птолемея (1984).

**3.** В соответствии с Талмудом (II в., см. Рабинович 1974, с. 352) за объем *стандартного* куриного яйца принималось среднее из объемов наибольшего и наименьшего из них (видимо из большого их числа).

**4.** А вот обобщающее мнение о Птолемее, написанное в более светлую пору советского периода (Чеботарев 1958, с. 579): его система “держала в духовном плену человечество в течение 14-ти веков” …

**5.** Возможно, что приближенные вычисления относились к *низшей*, прикладной, а не *научной* математике (Лурье 1934, с. 37).

**6.** И вот мнение о градусном измерении VIII в. в Китае (Нидем 1962, с. 51):

*По всей видимости, И-Син считал весьма нежелательным соглашаться … с сырой массой исходных данных, выказывающих значительный разброс, и, не будучи в состоянии статистически* *оценить их, использовал их лишь для того, чтобы удовлетвориться в том, что его вычисленные значения оказались примерно такими, какими они должны были быть. … И-Син вероятно полагал, что они были намного надежнее, чем большинство наблюдений.*

**7.** Этот подход соответствует качественному характеру древней науки. Так, древние географы выделили климатические пояса, хоть и не знали ничего об измерении температуры воздуха, а древние врачи (Гиппократ) полагали, что полные люди склонны умирать сравнительно рано, хотя никакой регистрации смертности не было. Оба эти примера иллюстрируют первую стадию статистического метода, когда из общих представлений делались качественные выводы (Шейнин 2005a, с. 9 – 10).

**8.** Численные данные очевидно подтверждают это предположение, поскольку спаренные наблюдения сильно отличались друг от друга. Так, первый одиночный результат был 0.″44, а составляющие первой пары наблюдений отличались от этого значения на 3′32″ и – 4′21″.

**9.** Мы видели немецкий перевод этого сочинения, опубликованный в Штутгарте без титульного листа или даты вместе с другими сочинениями Цицерона, каждое со своей пагинацией.

**10.** Ламберт (1765a, § 420) упомянул принцип минимакса, но добавил, что не сумел им воспользоваться общим методом и без многих окольных путей. См. также § 9.3.3.

**11.** В метеорологии крайние наблюдения возможно еще более ненадежны, чем в других отраслях науки, и средние из них поэтому более сомнительны. Добавим, что Якоб Бернулли, в своем *Дневнике* (частично опубликованном на том же латинском языке в 1975 г., см. с. 47, заметил, что при выводе среднего показания барометра следует отдавать предпочтение среднему арифметическому. И всё-таки даже в XIX в., несмотря на позднейшие соображения, метеорология всё еще использовала средние значения из крайних (Шейнин 1984b, с. 74).

**Основная литература**

Бируни (1967), Галилей (1632), Джинджерих (1983), Хальд (1990, гл. 10), Кеплер (1609), Нейгебауер ((1983), Птолемей (1998), Шейнин (1993b)

## Восемнадцатый век

Во второй половине этого века были достигнуты исключительно важные успехи и возникла сама теория ошибок. Ранние мемуары Лапласа мы, впрочем, рассматриваем в следующей главе.

**2.1. Среднее арифметическое.** Напомним (§ 1.9), что во времена Кеплера или несколько раньше среднее арифметическое видимо стало общепринятой оценкой "истинного значения" измеряемых констант (§ 6.6.1). Котс очевидно был первым, кто оставил явное утверждение о среднем, заявляя, что оно обеспечивает "наиболее вероятное" значение. Ни пояснения, ни примеров он не привел, однако его авторитет наверное подкрепил общее чувство. Лаплас (1814/1999, с. 862, левый столбец) заявил, что примеру Котса "следовали все вычислители", но описал его рекомендацию слишком расширенно, а до этого указал (1812, с. 352), что первым последователем Котса был Эйлер (1749). Вот, однако, выдержка из сочинения Котса (1722, с. 22):

> *Пусть p — положение какого-то предмета, определенное из первого наблюдения, q, r, s — его же положения из последующих наблюдений; пусть кроме того P, Q, R, S – веса, обратно пропорциональные расстояниям, на которые могут рассеиваться ошибки отдельных наблюдений и которые могут быть получены из данных о пределах ошибок. Будем считать, что веса ... соответствуют точкам положений p, q, r, s; найдем их* [положений] *центр тяжести Z. Я утверждаю, что точка Z будет наиболее вероятным положением предмета, которое с наивысшей вероятностью может считаться его истинным положением.*

Котс приложил и рисунок (быть может относящийся к трехмерному пространству), пометив на нем точки *p, q, r, s.*

Деламбр (посм. публ. 1827, с. 455) был осторожнее Лапласа:

> *Этот способ* [применение среднего] *употреблялся с наших дней многими геометрами, которые могли сами придти к нему, что не умаляет заслуги Котса, который видимо первым заимел эту мысль.*

Следующим ученым, который выразил свое мнение о среднем арифметическом, был Кондамин (1751, с. 223):

> *Если принять ... среднее из большого числа наблюдений, риск ошибки окажется невелик. И даже если среди этого большого числа некоторые наблюдения заметно ошибочны, средний результат едва изменится* [будет лишь немного искажен], *потому что избыток или недостаток в* [ошибочных] *наблюдениях распределится среди всех остальных и мало изменит результат.*

**2.2. Майер.** Он (1750) видимо был первым, кто решил систему линейных уравнений (1) после их разделения на группы и вычисления промежуточных частичных решений. Имея 27 уравнений с тремя неизвестными, он выделил 3 непересекающиеся группы по 9 уравнений в каждой и вычислил неизвестные, решив 3 суммарных уравнения. Пусть, к примеру, первая группа состоит из первых девяти уравнений, тогда Майер свел бы ее к единому уравнению, приняв, что

$$v_1 + v_2 + \ldots + v_9 = 0 \qquad (1)$$

(см. обозначения в § 1.1). И он справедливо заметил, что разбивка 27 уравнений на группы по 3, решение всех этих групп и осреднение полученных результатов было бы слишком трудным делом, ср. метод попарных сочетаний в §§ 2.5.1 и 4.1.

Майер в основном интересовался лишь одним неизвестным (назовем его $x$) и соответственно в первую группу он включил уравнения с наибольшими положительными коэффициентами $a_i$ при $x$, во вторую – уравнения с наибольшими отрицательными коэффициентами $a_i$. Заметим, что, решай он эти группы по МНКв, он получил бы значительный коэффициент [*aa*] и, соответственно, вес неизвестного $x$ оказался бы большим, см. формулу (5.14) и §5.3.5c.

Майер также оценил точность своей собственной оценки этого неизвестного, но принял при этом, что вес оценки возрастает вместе с числом уравнений (а не с корнем квадратным из этого числа).

В письме Шумахеру 24.6.1860 Гаусс (1865/1975, № 6, с. 90) заметил, что "Тоб. Майер вычислял не в соответствии с методическим принципом, а доморощенными сочетаниями". Он сослался на рукописи Майера, но возможно, что и в них метод вычислений был примерно тем же. Во всяком случае, Гаусс сам, в письме того же года (там же, с. 66 – 67) описал аналогичный метод, при помощи которого он, правда, градуировал анероид, а не устанавливал закономерность в природе[1].

**2.3. Отказ от уравнивания.** Уравнивание наблюдений до внедрения МНКв мы рассматриваем и ниже, но сейчас остановимся на особом обстоятельстве. Длительное время уравнивание триангуляции оставалось весьма затруднительным ввиду неоднозначности результатов вычислений. Мопертюи (1738, с. 160; 1756, с. 311 – 319) вычислил свою триангуляцию 12 раз (учитывая различные наборы измеренных углов), отобрал два варианта и выбрал из них среднее.

Но подчас астрономы XVIII в. не решались уравнивать триангуляции, опасаясь распространения крупных ошибок (Брю 1988, с. 225 – 226). Далее, в конце века Лаплас, Лежандр и другие ученые отказались уравнивать цепь триангуляции, проложенную между двумя базисами, и взамен решили вычислить каждую половину цепи от ближайшего к ней базиса (Шейнин 1993a, с. 50). Много позже Лаплас (прим. 1819, с. 590 – 591) обосновал этот отказ отсутствием в то время “истинной теории” уравнивания и добавил, что положение изменилось после того, как он обосновал МНКв. См также конец § 9.5.2.

**2.4. Ламберт.** Он был многосторонним ученым и свои сочинения (1760; 1765a; 1765b) частично посвятил обработке наблюдений. В первом из них он разделил ошибки по их происхождению (§ 282), доказывал, что крайние наблюдения следует отбрасывать (§§ 287–291) и оценивал точность наблюдений (§ 294), а далее (§ 303) сформулировал принцип наибольшего правдоподобия. Мы опишем все указанные сочинения; наши §§ 2.4.1 – 2.4.3 относятся к первому из них и именно к его первоначальному латинскому изданию, поскольку указанное выше было выпущено из немецкого перевода как устаревшее ...

**2.4.1.** Отбраковку отклоняющихся наблюдений Ламберт пытался обосновать тем, что они искажены наибольшими по абсолютной величине ошибками. Это утверждение бесполезно, поскольку нам становятся известны не ошибки, а уклонения от среднего. Ламберт, правда, добавил, что имел в виду наблюдения, значительно уклоняющиеся от остальных.

**2.4.2.** Оценка точности. Пусть $\bar{u}$ – среднее арифметическое из $n$ наблюдений и $\bar{v}$ – такое же среднее из ($n$ – 1) наблюдений, оставшихся после исключения одного из них. Тогда, как заявил, но не до-

казал Ламберт, $\bar{u}$ вряд ли уклоняется от искомого истинного значения измеряемой константы более, чем на $|\bar{u} - \bar{v}|$. Если отброшенное наблюдение было грубо ошибочно, то оценка Ламберта окажется слишком неудачной, и в любом случае она не нормирована (не зависит от количества наблюдений). Полезнее было бы применить несколько значений $|\bar{u} - \bar{v}|$, полученных при $n$,
$(n - 1)$, $(n - 2)$, и т. д.

**2.4.3.** Принцип наибольшего правдоподобия. Ламберт (§ 295) указал, что будет отыскивать среднее, вероятность наименьшего отклонения которого от истинного значения окажется наивысшей. От истинного значения отошла только математическая статистика, в теории ошибок оно осталось, но получило определение (§ 6.6.1). По существу же утверждение Ламберта представляется неверным: наименьшее уклонение просто равно нулю. **Здесь рисунок**

Ламберт далее нарисовал некую непрерывную одновершинную плотность распределения, общий вид которой соответствовал свойствам "обычных" случайных ошибок. Ординаты *PN, QM, …* этой кривой он назвал "истинными количествами появлений" соответствующих ошибок *CP, CQ, …* (C – мода кривой), т. е. применил понятие, свойственное дискретному случаю. Он предположил, что эти ошибки фактически произошли *n, m, …* раз, ввел условие

$$PN^n QM^m \ldots = \max \qquad (2)$$

и пояснил определение максимума геометрическим путем (при помощи подкасательных). Показатели степени $n,\ m,\ \ldots$ вряд ли были нужны, поскольку они исчезли бы при бесконечно малом изменении соответствующих результатов наблюдений.

Обозначим уравнение плотности через $\varphi(x - \hat{x})$, где $\hat{x}$ –неизвестная мода, а наблюдения через $x_1, x_2, \ldots, x_n$, тогда уравнение (2), но без указанных показателей степени, примет вид

$$\varphi(x_1 - \hat{x})\,\varphi(x_2 - \hat{x})\ldots\,\varphi(x_n - \hat{x}) = \max.$$

Впрочем, Ламберт (§ 306) полагал, что оценка наибольшего правдоподобия обычно совпадала со средним арифметическим.

**2.4.4.** Среднее арифметическое. Ламберт (1765a, § 320) назвал его "наверняка самым надежным", если только положительные и отрицательные ошибки равновозможны, и добавил (1765b, § 3), что при возрастании числа наблюдений оно стремится к искомому истинному значению. Мы вернемся к этому и аналогичным утверждениям 1760 г. в § 6.6.1.

Ламберт (1765a, § 441) также заметил, что применение среднего основано на его наибольшей вероятности, что, конечно, верно только если оно совпадает с модой. Все эти утверждения он основал на сравнении (1765a, §§ 443 – 445) среднего арифметического со средним из крайних наблюдений и они во всяком случае не были достаточно понятны.

**2.4.5.** Кривая плотности. Ламберт (1765a) снова разделил ошибки по их происхождению (§311), затем проверил свойства ошибок элементарным графическим приемом (§§ 435 – 436) и вывел кривую плотности ошибок наведения геодезического инструмента на цель. **рисунок**

Вот этот вывод. Наводя трубу инструмента на некоторую точку *C*, наблюдатель фактически фиксирует любую точку на интервале *AB,* серединой которого является *C*. Возможность наведения на некоторую точку *F* на *AB* равна длине отрезка *DE*, перпендикулярного *AB* и проходящего через *F* таким образом, что *DF* = *EF*. Крайние точки подобных перпендикуляров образуют замкнутую кривую, симметричную относительно *AB,* и именно окружность, поскольку "нет причин" для "угловатости", ср. § 3.1.2. Формально вывод Ламберта означал, что вероятность ошибки $x$ равнялась

$$\varphi(x) = (2/\pi r^2)\sqrt{r^2 - x^2},\ |x| \le r,\ r \text{ неизвестно.}$$

**2.4.6.** Уравнивание косвенных наблюдений. Ламберт (1765b, см., например, § 24) подбирал прямую к множеству точек наблюдений $M_i(x_i; y_i)$. Он делил эти точки на две группы (с меньшими и большими абсциссами), определял центр тяжести каждой группы и проводил прямую через них. Если уравнение прямой записать в виде

$$ax + by + c = 0, \qquad (3)$$

то его параметры можно будет определять по координатам центров тяжести 1 и 2:

$$a\bar{x}_1 + b\bar{y}_1 + c = 0,\ a\bar{x}_2 + b\bar{y}_2 + c = 0. \qquad (4a;\ b)$$

В то же время

$$ax_i + b_i y + c_i = v_i,$$

откуда следует, что уравнения (4a) и (4b) могут быть получены из условий равенства нулю соответствующих сумм $\sum v_i$, т. е. из условий, которые применял и Майер (§ 2.2).

Ламберт (1765b, § 22) также указал на возможность построения дополнительной прямой по $(n - 1)$ точкам, а расхождение между двумя прямыми послужило бы, по его мнению, мерой надежности результата, ср. § 2.4.7[2].

**2.4.7.** Теория ошибок. Этот термин (*Theorie der Fehler*) ввел Ламберт (1765a, Vorberichte). Там же, в § 321, он определил ее задачи: установление соотношений между ошибками, их последствиями (Folgen), обстоятельствами наблюдения и надежностью инструмента. Ни Лаплас, ни Гаусс ни разу не употребили этот новый термин, но вот Бессель (1820, с. 166; 1838b, § 9) либо перенял его, либо ввел его независимо. В середине века этот термин вошел в употребление также на французском и английском языках (Fischer 1845, заглавие его Abschnitt 1; Лиагр 1852, заглавие; Эйри 1861, заглавие)[3].

**2.5. Бошкович**

**2.5.1.** В 1750 – 1753 гг. он вместе с другим астрономом, Мейром (Maire) измерили длину градуса в Италии, а затем Бошкович определил параметры земного эллипсоида вращения по уравниванию результатов нескольких ученых. При этом он иногда применял своеобразный метод вычисления среднего арифметического (Чубранич 1961, с. 46). Имея наблюдения зенитных расстояний тех же самых двух звезд на обоих концах своей дуги, в Риме и Римини, притом два наблюдения в Риме и одно – в Римини, Бошкович вычислил 2 значения разности широт $\Delta\varphi$ этих городов по каждой из этих звезд и молчаливо назначил всем четырем разностям один и тот же вес. Он далее вычислил все 6 полусумм слагаемых $\Delta\varphi_i$ и $\Delta\varphi_j$, $i, j = 1, 2, 3, 4$, $i < j$, и среднее из них. Такой косвенный путь быть может применялся для качественной оценки расхождений наблюдений и выявления менее надежных из них[4].

Первоначальный метод Бошковича уравнивания косвенных наблюдений (Мейр и Бошкович 1770, с. 483 – 484; Чубранич 1961, с. 90 – 91) был в этом отношении аналогичен. Он решал системы с двумя неизвестными (упомянутыми выше параметрами) методом попарных сочетаний, – составляя уравнения попарно, решая каждую пару, пренебрегая величинами $v_i$ и осредняя полученные частные решения, ср. §4.1.

Сравнив результаты 10 парных сочетаний пяти градусных измерений, Бошкович решил, что нужен новый способ уравнивания. Среднее, как он (Мейр и Бошкович 1770, с. 501) убеждал, не должно быть "простым средним арифметическим", оно обязано быть "связано определенным законом с правилом случайных сочетаний и исчислением вероятностей". Соответственно, он посчитал, что $v_i$ должны быть подчинены условиям

$$v_1 + v_2 + \ldots + v_n = 0, |v_1| + |v_2| + \ldots + |v_n| = \min. \qquad (5; 6)$$

Первое из них, продолжал Бошкович, следует из равной вероятности ошибок каждого знака, второе же "необходимо, чтобы как можно теснее приблизиться к наблюдениям ..." И всё же, не применяя не известных еще идей и методов, относящихся к плотностям распределения, он не смог сказать, как именно его условия были *связаны* с вероятностными правилами.

От условия (5) легко освободиться, суммируя все исходные уравнения и исключая одно из неизвестных, так что существенно лишь второе условие. Позднее Гаусс (§5.1.1) заметил, что оно приводит к определенному числу нулевых $v_i$ и, соответственно, не согласился с этим методом. Лаплас (§ 3.6.4), однако, решал систему уравнений по методу Бошковича (в частности, в т. 2 своей *Небесной механики*), а У. Гершель (1805), см. также Шейнин (1984а, с. 172 – 173), фактически применил условие (6), которое, напомним, впервые появилось у Галилея (§ 1.7).

В случае градусных измерений исходные уравнения принимают вид

$$a_i x + y + l_i = v_i, \; i = 1, 2, \ldots, n. \qquad (7)$$

Если сложить их и разделить сумму на $n$, то ввиду условия (5)

$$\bar{a}x + y + \bar{l} = 0.$$

Вычитая это из каждого из уравнений (7), получим

$$(a_i - \bar{a})x + (l_i - \bar{l}) = v_i. \quad (8)$$

Далее, как будто бы представляется, что

$$x = -(l_i - \bar{l})/(a_i - \bar{a}),\ i = 1, 2, \ldots, n, \quad (9)$$

но условие (6) определяет медиану непосредственных наблюдений, так что искомое значение неизвестного должно быть медианой чисел (9).

При переводе *Небесной механики* на английский язык Боудитч (Лаплас 1798 – 1825/1832, § 40, примечание) заметил по этому поводу:

> *Метод наименьших квадратов, при его применении к системе наблюдений, в которых одна из крайних ошибок очень велика, обычно не приводит к столь верному результату, как метод, предложенный Бошковичем. ... Причина здесь в том, что в первом методе крайняя ошибка* [как и всякая другая] *влияет на результат пропорционально своему квадрату, а во втором – пропорционально первой степени ...*

Метод Бошковича устойчив, потому что связан с медианой.

**2.5.2.** Случайные суммы. В рукописи без даты Бошкович чисто комбинаторно вычислил шансы ошибки в сумме $n$ ошибок наблюдений, молчаливо полагая их взаимно независимыми и указывая, что каждая из них с равными вероятностями принимает значения – 1, 0 и 1. Он закончил свое исследование случаем $n = 8$, не упомянув ни перехода к большему значению $n$, ни более общего дискретного равномерного распределения, ни даже среднего арифметического. К тому же, он пояснил формулу для подсчета числа сочетаний (и ввел для него свое собственное и малопригодное обозначение), так что его рукопись, видимо, не предназначалась ни математикам, ни астрономам.

Бошкович имел предшественников: Якоба Бернулли, 1713, Галилея (посм. публ. 1718 г.) и Лейбница (рукопись 1676 г, опубликованная Бирманом в 1956 г.). Все они исследовали ту же задачу в терминах игры в кости.

В другом месте Бошкович (1758, § 481, см. также § 479) описывал движение частиц, "движущихся совместно практически с одной и той же скоростью" и несколько туманно указал, что сумма (не среднее!) $n$ "иррегулярных неравенств" между скоростями стремится к нулю при неограниченном возрастании $n$.

Это неверное заключение (ср. мнение Кеплера о весе суммы монет в § 1.9) возможно было навеяно описанными выше рассуж-дениями[5]. Если это так, то рукопись Бошковича о поведении ошибок наблюдения была написана до мемуара Симпсона 1756 г. (§ 2.6).

Заметим еще, что он всё-таки не считал, что "частицы" материи движутся с одной и той же скоростью.

**2.6. Симпсон.** В Новое время длинные серии наблюдений выполнили и Браге (§ 1.8), и позднейшие астрономы, в первую очередь Брадлей. Описывая свое открытие нутации, Брадлей (1750, с. 17) заметил:

*Это* [открытие] *показывает нам громадную пользу совершенствования* [астрономии], *равно как и всякой иной ветви естественных наук регулярными рядами наблюдений и опытов*[6].

Тем не менее, некоторые естествоиспытатели, хотя вряд ли астрономы, предположили, что один-единственный опыт (не наблюдение!) может оказаться ценнее, чем их множество (Бойль, посм. публ. 1772/1999, с. 376). Возможно, что это разумное замечание иногда недопустимо обобщалось. Во всяком случае, Симпсон (1756/2006, с. 116) попытался опровергнуть

*Некоторых весьма известных лиц, полагавших и даже публично заявлявших, что одному-единственному наблюдению, выполненному с должной тщательностью, следует доверять так же, как среднему из очень большого их числа.*

Даже если Бошкович (§ 2.5.2) предвосхитил Симпсона в вероятностном рассмотрении ошибок наблюдения, что вовсе не очевидно, именно Симпсон опубликовал первое и притом действительно важное исследование по этой теме. Без всяких пояснений он принял (еще до Ламберта, см. § 2.4), что погрешности обладают плотностью распределения, – дискретной равномерной, а затем (второй случай) – дискретным треугольным, а в целом его мемуар означал, что Симпсон рассматривал ошибку наблюдений как случайную величину, которую лишь Пуассон (1837, с. 140 – 141) определил формально (и назвал ее явно временным термином, – вещью *A*). Васильев (1885, с. 127 – 131) обсуждал случайные величины, не объясняя этого термина, а на с. 133 заметил, что случайные ошибки обладают всеми их свойствами, – и своими собственными.

**2.6.1.** Равномерное распределение. Значения ошибок

$$-v,\ -(v+1),\ \ldots\ -1,\ 0,\ 1,\ \ldots,\ (v-1),\ v \qquad (10)$$

равновероятны. Обозначим ошибку наблюдения $i$ через $\varepsilon_i$, а число некоторых шансов через $N$. Тогда, как заметил Симпсон,

$$N = N(\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n = m)$$

есть коэффициент при $r^m$ в разложении

$$(r^{-v} + \ldots + r^0 + \ldots + r^v)^n = r^{-vn}\,(1-r)^{-n}\,(1-r^{2v+1})^n$$

и поэтому

$$N = C_{n+q-1}^{q} - C_n^1 \; C_{n+q-w-1}^{q-w} + C_n^2 \; C_{n+q-2w-1}^{q-2w} - \ldots =$$
$$C_{n+q-1}^{n-1} - C_n^1 \; C_{n+q-w-1}^{n-1} + C_n^2 \; C_{n+q-2w-1}^{n-1} \qquad (11)$$

Здесь $q = nv + m$, $w = 2v + 1$ и ряд продолжается пока биномиальные коэффициенты еще имеют смысл. Соответствующая вероятность равна, разумеется, $N/(2v+1)^n$.

**2.6.2.** Треугольное распределение. Вероятности ошибок (10) здесь пропорциональны

$$1, 2, \ldots (v-1), v, (v-1), \ldots, 2, 1.$$

Аналогичным образом Симпсон заметил, что

$$N = N(\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t = m)$$

есть коэффициент при $r^m$ в разложении

$$[r^{-v} + 2r^{-v+1} + \ldots + (v+1)\,r^0 + \ldots + 2r^{v-1} + r^v]^t =$$
$$r^{-vt}\,(1-r)^{-2t}\,(1-r^{v+1})^{2t}$$

и поэтому

$$N = C_{p-1}^{n-1} - C_n^1 \; C_{p-w}^{n-1} + C_n^2 \; C_{p-2w}^{n-1} - \ldots \qquad (12)$$

Здесь $n = 2t$, $p = tv + m + n$ и $w = v + 1$.

Для этого распределения Симпсон также вывел $P[\sum(\varepsilon_i/t) = m/t]$ и $P[(\sum\varepsilon_i/t) \leq m/t]$ и во втором случае сослался на "метод приращений", т. е. на суммирование обобщенных степеней вида

$$x(x-1)(x-2)\ldots(x-k+1).$$

Итак, Симпсон по существу применил производящие функции вида

$$f(r) = \delta_{-v}r^{-v} + \delta_{-v+1}r^{-v+1} + \ldots + \delta_0 r^0 + \ldots + \delta_{v-1}r^{v-1} + \delta_v r^v,$$

полагая, соответственно, для рассмотренных распределений

$$\delta_k = 1 \text{ и } \delta_{-v} = \delta_v = 1,\ \delta_{-v+1} = \delta_{v-1} = 2 \text{ и т. д.}$$

Для обоих распределений Симпсон определил вероятность абсолютной погрешности среднего арифметического быть меньше некоторой величины или равняться ей и решил, что это среднее вообще (стохастически) предпочтительней отдельного наблюдения и тем самым произвольно и неверно расширил доказанное им.

Формулы типа (11) при помощи производящих функций выводил Муавр (1712/1984, с. 240, без доказательства; 1730, с. 191 – 197; посм. публ. 1756 г., с. 39 – 41), который, однако, не рассматривал ошибок наблюдений. Сам Симпсон (1740, Задача 22) раннее исследовал игру в кости, вполне аналогичную своему нынешнему случаю равномерного распределения. И теперь, в 1756 г., он, естественно, заметил, что его формула (11) также определяет число шансов для появления $(n + q)$ очков при броске $n$ костей с $w$ гранями каждая и что формула (12) соответствует появлению $p$ очков при том же броске.

Второй случай Симпсона более интересен, потому что, в отличие от первого, он, правда, грубо соответствует действительности.

**2.6.3.** Годом позже Симпсон обобщил свое исследование, рассмотрев непрерывное треугольное распределение, – первое непрерывное распределение ошибок наблюдения, хотя и не первое в теории вероятностей. Действительно, в 1733 г. Муавр ввел нормальное распределение, а еще в 1709 г. Николай Бернулли (Тодхантер 1865, с. 195 – 196) рассмотрел непрерывный равномерный закон смертности.

Симпсон также нарисовал график плотности для средней ошибки, – первый график в теории ошибок, но опять-таки не в теории вероятностей. Еще в 1669 г. Гюйгенс, в посм. публ. 1895 г., между с. 530 и 531, см. также Шейнин (2005а, с. 41 – 42), нарисовал непрерывную кривую, описывающую смертность, уравнение которой можно сейчас представить в виде $y = 1 - F(x)$.

В своем непрерывном случае Симпсон исходил из формулы (12), предполагая, что $|v| \to \infty$ при постоянном соотношении $(m/n)/v$, где дробь в числителе есть допустимая погрешность среднего арифметического и $n$ по-прежнему число наблюдений. Он получил

$$P[(\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t)/t \le m/t] =$$
$$1 - (2/n!)\{(p/v)^n - n[(p/v) - 1]^n + C_n^2\ [(p/v) - 2]^n - \ldots\}.$$

Его объяснения были недостаточны. Фактически он посчитал, что ошибки (10) становятся равными $-kv, -k(v-1), \ldots$ и что $v \to$ и $k \to 0$ при $kv = h$. В то же время $m$ становится равным $mkv = mh$ и остается конечным, а не бесконечным, как утверждает Симпсон. Наконец, возможно полагать, что $h = 1$ (Тодхантер 1865, с. 308).

График Симпсона соответствовал конечному $v$ при непрерывном аргументе (погрешности наблюдения), а кривая ошибок среднего арифметического не имела характерного для нормального распределения закругления. Симпсон, разумеется, не владел понятием дисперсии и подсчет вероятности того, что абсолютная погрешность среднего превзойдет ту же погрешность одного наблюдения оказался нелегким (Шусмит 1985).

**2.7. Лагранж.** Он опубликовал длинный мемуар (1776) о вероятности сумм и средних значений ошибок наблюдений. Будучи математиком, а не естествоиспытателем, он рассмотрел ряд дискретных и непрерывных распределений, явно не имевших отношения к своей объявленной цели, в том числе, например, закон косинуса. Не упоминая Симпсона[7], он включил и оба его распределения, а в своем § 18 впервые упомянул *кривую возможностей ошибок*.

Карл Пирсон (посм. публ. 1978, с. 587 – 612) подробно описал мемуар Лагранжа и отметил в нем интересные общематематические нововведения, а на с. 599 указал, что в своем § 6 Лагранж по существу оценил член полинома по "теореме Стерлинга" и пришел к многомерной нормальной поверхности[8].

**2.8. Даниил Бернулли.** В теории вероятностей он и Муавр были двумя наиболее влиятельными предшественниками Лапласа. К нашей теме относятся два его мемуара.

**2.8.1.** Иоганн III Бернулли (1789) описал рукопись Даниила, которую он получил в 1769 г. и которая была недавно опубликована в переводе (1997). В ней Даниил принял плотность распределения ошибок наблюдения $x_1$, $x_2$, ..., $x_n$ в виде кривой второго порядка, а за оценку $\hat{x}$ истинного значения измеряемой константы – обобщенное среднее арифметическое

$$\hat{x} = \sum p_i x_i / \sum p_i \qquad (13)$$

причем в случае параболического закона апостериорные веса оказались равными

$$p_i = r^2 - (\hat{x} - x_i)^2 \qquad (14)$$

и значение $r$ определялось разумным выбором наибольшей возможной ошибки для данного наблюдателя, которая, однако, вполне могла возрастать с числом наблюдений. Вычисление $\hat{x}$ было возможно методом последовательных приближений с начальным значением этой оценки, равным $\bar{x}$.

Первыми, кто применил подобную оценку, были Браге (§ 1.8) и Кеплер (§ 1.9), затем Шорт (1763). Следует однако добавить, что введение апостериорных весов сводится лишь к поправке обычного среднего арифметического за асимметрию фактического распределения, хотя, конечно же, дисперсия этого среднего меняется.

**2.8.2.** Позднее Даниил Бернулли (1778) опубликовал переработанный вариант своей рукописи. В нем он отрицал среднее арифметическое, при выборе которого "самый искусный лучник не имел бы никакого преимущества перед слепым" (§ 5). Малые ошибки, продолжал он, более вероятны, чем крупные[9], так что плотность можно принять в виде "полуэллипса", полуокружности, или, как он в конце концов решил, в виде дуги параболы

$$y = r^2 - (\hat{x} - x)^2 \qquad (15)$$

где $\hat{x}$ – искомая оценка, а $r$ имело то же значение, что и в рукописи.

Он далее предложил в качестве $\hat{x}$ оценку наибольшего правдоподобия, т. е. моду одновершинной кривой (15). Требуемые при этом вычисления оказались, однако, слишком тяжелыми. Действительно, условие

$$z = [r^2 - (\hat{x} - x_1)^2]\,[r^2 - (\hat{x} - x_2)^2]\,\ldots\,[r^2 - (\hat{x} - x_n)^2] = \max \quad (16)$$

приводит к

$$(p_1 + p_2 + \ldots p_n)\,\hat{x} = p_1x_1 + p_2x_2 + \ldots + p_nx_n, \qquad (17)$$

$$p_i = [r^2 - (\hat{x} - x_1)^2]^{-1}. \qquad (18)$$

Уравнение (17) можно, конечно, решить последовательными приближениями, но Бернулли этого почему-то не упомянул. Он подкрепил свое предложение некоторыми соображениями и численными примерами для $n = 3$ и приближенно решил для этого случая алгебраическое уравнение пятой (!) степени[10].

Веса (18) возрастают к краям распределения (14), о чем Бернулли умолчал, быть может потому, что такое поведение показалось бы странным и неприемлемым, к тому же оно противоречило и его прежнему выбору апостериорных весов, см. наш § 2.8.1, и его соображению о лучнике в § 5, см. выше. Что апостериорные веса подобного рода действительно могут применяться, было установлено лишь сравнительно недавно, см., например, Сархан и Гринберг (1962). Мы, впрочем, имеем в виду не относительные веса (14), а те, которые соответствовали бы кривой (15) после нормирования.

**2.8.3.** В своем втором мемуаре по теории ошибок Бернулли (1780) исследовал ошибки маятниковых наблюдений. Пусть из $2N$ суточных колебаний маятника ($2N \approx 86\ 400$), $(N + \mu)$ замедлены и имеют период $(1 + \alpha)$ сек, а $(N + \mu)$ ускорены и имеют период $(1 - \alpha)$ сек. Общая суточная ошибка маятника будет равна

$$\delta = 2N - (N + \mu)(1 + \alpha) - (N - \mu)(1 - \alpha) = -2\mu\alpha \qquad (19)$$

и, если $\mu = 100$ и $\alpha = 0.01$ *сек*, $\delta = 2$ *сек*.

Интереснее был следующий шаг. В одном из своих прежних мемуаров Бернулли (1770 – 1771) рассмотрел важную задачу из статистики населения. Именно, он оценивал относительные количества мужских и женских рождений и мог бы до Лапласа, примени он не суммирование, а интегрировавание, придти к теореме Муавра-Лапласа, правда после Муавра, на которого он не сослался .

Предположив, что из $2N$ младенцев ($N = 10\ 000$) $m$ мальчиков, а вероятности рождения обоих полов совпадают (во второй части мемуара он отказался от этого допущения), Бернулли доказал[11], что

$$P(m = N \pm \mu) \approx c \exp(-\mu^2/N)$$

и заметил, что вероятность неравенств $0 \le \mu \le 47$ равна 1/2.
Теперь, в 1780 г., $N = 43\ 200$ и для той же половинной вероятности он получил $\mu = 100$, указав также, что погрешность типа $\delta$, см. формулу (19), окажется равной $\delta/\sqrt{365}$ и $\delta/\sqrt{24}$ для года и часа соответственно. Обе эти оценки являлись простым следствием нормального закона, но после Гаусса (§ 5.3.4) стало известно, что они

(выраженные в терминах средней квадратической ошибки) имели бы место и в более общем случае.

Наконец, Бернулли отделил систематические (*хронические*) ошибки, чье влияние почти постоянно, от случайных (*моментных*), действующих пропорционально корню квадратному из соответ-ствующего промежутка времени. О своих представлениях о сути случайных ошибок Бернулли не сообщил.

На основе второй части своего мемуара 1770 – 1771 гг. Бернулли мог бы обобщить свои результаты на неравные числа замедленных и ускоренных колебаний, мог бы что-то сказать и о возможной зависимости смежных колебаний друг от друга, но не сделал этого. Он, однако, первым применил в теории ошибок нормальный закон и вероятную ошибку (формально введенную Бесселем в 1816 г.)[12] и первым явно выделил два вида ошибок[13].

**2.9. Эйлер.** Его основной вклад в обработку наблюдений это комментарий 1778 г. к первому мемуару Бернулли. Он (§ 2) указал, что апостериорные веса должны быть равными (14), а в § 4 ошибочно заметил, что Бернулли именно это и предложил. Ослепший в 1771 г., он, видимо, исходил здесь из общих соображений автора о лучнике (§ 2.8.2).

Далее, Эйлер (§ 6) опровергал принцип наибольшего правдоподобия, заметив, что при наличии наблюдения, которое "может быть либо отброшено, либо сохранено", произведение (16) даже в максимуме "окажется сведенным на нет", тогда как "принципы искусства предположений" (еще не теории вероятностей!) требуют, чтобы подобное наблюдение ни в том, ни в другом случае никак не воздействовало на результат. Вторая половина этого утверждения непонятна; см. также соответствующее мнение Гаусса (§ 5.3.2).

Исходя из уравнений (13) с весами (14), а не (18), Эйлер вывел кубическое уравнение с неизвестным $\hat{x}$. Он обозначил наблюдения через $\Pi + a$, $\Pi + b$, $\Pi + c$, ...

$$a + b + c \ldots = 0 \qquad (20)$$

и при $B = a^2 + b^2 + c^2 + \ldots$, $C = a^3 + b^3 + c^3 + \ldots$ получил

$$n\hat{x}^3 - \hat{x}\,(nr^2 - 3B) - C = 0, \qquad (21)$$

выразившись по поводу параметра $r$ примерно так же, как Бернулли. За $\hat{x}$, как он указал, следует принять наименьший по абсолютной величине корень этого уравнения, видимо потому, что ввиду условия (20) среднее арифметическое из $a, b, c, \ldots$ было равно нулю.

Наконец, Эйлер (§ 11) заметил, что его кубическое уравнение может быть получено из условия

$$[r^2 - (\hat{x} - a)^2]^2 + [r^2 - (\hat{x} - b)^2]^2 + [r^2 - (\hat{x} - c)^2]^2 + \ldots = \max, \quad (22)$$

т. е. из максимума суммы квадратов весов (14). Именно Бернулли

(§ 2) употребил выражение "веса или моменты", Эйлер же (§ 7) упомянул степени доброкачественности (ср. выше), – *pretium seu gradum bonitatis.*

Сумма в (22) равна трехчлену

$$nr^4 - 2r^2[(\hat{x} - a)^2 + (\hat{x} - b)^2 + (\hat{x} - c)^2 + \ldots] + [(\hat{x} - a)^4 + (\hat{x} - b)^4 + (\hat{x} - c)^4 + \ldots],$$

последний член которого сравнительно мал, так что условие (22) можно записать в виде

$$(\hat{x} - a)^2 + (\hat{x} - b)^2 + (\hat{x} - c)^2 + \ldots = \min,$$

что является принципом наименьших квадратов, хотя только для случая одного неизвестного.

Не скажи Эйлер "степени доброкачественности", можно было бы считать, что просто принципом среднего арифметического, но следует оговориться еще раз: и Эйлер, и Бернулли пытались отыскать что-то лучшее чем среднее арифметическое, так что первый подход Гаусса 1809 г. (§ 5.1.3) здесь невозможен. Далее, приняв для ошибок наблюдений определенную плотность, они тем самым заявили, что, вообще говоря, возможно улучшить оценки МНКв, т. е. второго гауссова обоснования 1823 г., которое не было связано ни с каким распределением.

Бернулли не мог не заметить ошибки Эйлера в выборе весов, но, видимо, смолчал. Упомянем еще, что оба они в свое время читали *Фотометрию* Ламберта, см. Бопп (1924, с. 15 – 17) и Раделе-Де Граве и др. (1979, с. 73 – 74), в котором впервые был предложен принцип наибольшего правдоподобия (§ 2.4), но не сослались на нее.

**Примечания**

**1.** Приближенные вычисления, конечно же, не вышли из употребления. Бессель (1826, с. 229) при исследовании термометров заключил, что 26 исходных (не нормальных) уравнений с таким же числом неизвестных было бы слишком трудно решать по МНКв и применил приближенный метод.

**2.** Тиллинг (1975, с. 201 – 206) описала использование графиков Ламбертом.

**3.** В письме 1971 г. ко мне покойный Э. Ш. Пирсон пояснил, почему его отец не описал работ Ламберта в своей посмертной и позднее вышедшей под его редакцией книге (Пирсон 1978):

> *Это произошло не потому, что сочинения* [Ламберта] *были написаны по-немецки, которым мой отец отлично владел. Я полагаю, … что он выбрал для изучения тех ученых, которые были указаны в небольшом числе источников, например в трактате Тодхантера, и что эти источники не включали имя Ламберта. Конечно, ко времени, к которому его лекции перешли за 1750-й год, К. П. было уже за 70, и его исследование безусловно ограничивалось четырьмя французами, – Кондорсе, Даламбером, Лагранжем и Лапласом.*

Тодхантер (1865) всё-таки упоминал Ламберта, но не описал его трудов.

**4.** В более интересных случаях, по крайней мере в XIX в., каждое наблюдение делалось *своим* прибором. Так, при измерении относительного ускорения силы тяжести на одной из своих станций Сабин (1821) получил 4 результата, назовем их $x_{11}$, $x_{12}$, $x_{21}$, и $x_{22}$, каждое из которых соответствовало определенному сочетанию его часов и маятников; $x_{21}$, к примеру, относилось к часам № 2 и маятнику № 1.

**5.** Даже много позже Гельмерт (1905, с. 604) счел нужным указать, что сумма случайных ошибок вовсе не стремится к нулю.

**6.** Вот интересное замечание Декарта (1637/1982, с. 63):

*По поводу экспериментов я замечаю, что чем дальше мы продвигаемся в наших знаниях, тем необходимее они становятся.*

**7.** Возможная, хотя и недостаточная причина состояла в ожесточенном споре о приоритете между Муавром и Симпсоном и Лагранж видимо не хотел даже косвенно вмешиваться в нее. Муавр был гораздо более значимым ученым, нежели Симпсон, притом старше его на 43 года. В нескольких существенных случаях Симпсон не сослался на Муавра, и, будучи обвинен последним (Муавр 1743, с. XII; в позднейших изданиях отсутствует) "в показе моих новых правил и работ", воззвал "ко всему человечеству" с риторическим вопросом (Симпсон, посм. публ. 1775, с. 144), не выказал ли Муавр "самонадеянность, дурной характер и закоренелость, не подобающие джентльмену".

**8.** В этом мемуаре Лагранж смог воспользоваться производящими функциями в непрерывном случае и таким образом проложить путь характеристическим функциям и опубликовал *первый список преобразований Лапласа* (Сил 1949/1977, с. 72; Шейнин 1973а, §2).

**9.** Карл Пирсон (посм. публ. 1978, с. 268), удачно заметил, что небольшие ошибки в таких случаях происходят чаще и потому "внесут свой надлежащий вклад в среднее арифметическое", от которого, стало быть, не следует отказываться.

**10.** Значения $x_1$, $x_2$, … в подобных примерах следует выбирать в соответствии с принимаемой плотностью. Впрочем, для случая трех наблюдений это не столь важно.

**11.** Даниил Бернулли вывел лишь локальную теорему; в нескольких случаях он суммировал вероятности, но интегралов не вычислял.

**12.** В 1669 г. Людовик и Христиан Гюйгенс ввели в своей переписке вероятную продолжительность жизни (Гюйгенс 1669/1895, с. 531 – 532 и 537).

**13.** Уже Птолемей (§ 1.4) имел некоторое представление об этом. И вот мнение Уайтсайда (1972), издателя сочинений Ньютона, из его письма ко мне в связи с публикацией Шейнин (1971а) о предосторожностях, которые Ньютон принимал в своих опытах:

*Фактически (но без явных утверждений об этом) Ньютон четко представлял себе различие между случайными и структурно <u>встроенными</u> ошибками. Он безусловно был погружен в мысли о втором типе <u>встроенных</u> ошибок и многие теоретические модели различных видов физических, оптических и астрономических явлений были сознательно продуманы им таким образом, чтобы* *свести к минимуму эти структурные ошибки. В то же время он подходяще регулировал свою практическую астрономическую работу в смысле случайных ошибок наблюдений.*

Мы не сомневаемся в том, что подобное мнение можно было бы высказать, например, по поводу Гюйгенса.

**Основная литература**

Бернулли Д. (1778; 1780), Ламберт (1760; 1765a; 1765b), Майер (1750), Пирсон К. (1978), Симпсон (1757), Стиглер (1986), Чубранич (1961), Шейнин (1971b; 1972a; 1972b; 1973a; 1973b; 1973c; 1975; 1993b), Шейнин и Майстров (1972)

## 3. П. С. Лаплас

### 3.1. Введение

**3.1.1.** Лаплас был астрономом и физиком и внес впечатляющий вклад в математику, но математиком себя не чувствовал. Так, введя интегралы от функций комплексного переменного, он (1810a, с. 304) заметил, что надеется, что *геометры* заинтересуются этим. В теории вероятностей Лаплас не предложил даже эвристического определения случайной величины и не рассматривал ни плотности распределения, ни характеристические функции (последние он сам же и ввел), как самостоятельные математические объекты. Применив, например, различные формулы для определения плотности случайных сумм, он (Шейнин 1973a, § 3) не выписал ни одной из них, а лишь решил несколько соответствующих и обособленных задач. Уровень абстракции в его трудах оказался недостаточно высоким, и после него теорию вероятностей пришлось создавать заново.

Лаплас не смог бы добиться столь выдающихся успехов в естественных науках без предварительного продвижения теории ошибок, но в его исполнении она также не выдержала испытания временем (§ 5.4.4).

Наши §§ 3.2 – 3.7 посвящены отдельным сочинениям Лапласа, а в § 3.8 мы обсуждаем один специальный вопрос.

**3.1.2.** Несколько раз в своих ранних мемуарах Лаплас (1774, см. наш § 3.2.1a; 1776, с. 148; 1781, см. наш § 3.2.2) выводил законы распределения ошибок наблюдения исходя из незнания. Его *классическое* определение вероятности (которое, однако, восходит к Якобу Бернулли) было не лучше. Впрочем, оно удержалось до наших дней и было лишь дополнено непризнанным определением Мизеса и *теоретически* заменено в аксиоматике. Наконец, Лаплас (1776, с. 144 – 145) даже заявил, что именно незнание причин привело к возникновению теории вероятностей; на самом же деле –необходимость выявлять закономерности, присущие массовым случайным явлениям.

Некоторые философы науки отрицают за незнанием причин разумную отправную точку и возражают против "принципа неопределенности" (Кейнс 1921/1973, с. 44), или, как он раньше назывался, "принципа недостаточных оснований" (Крис 1886, с. 6). Лаплас (1798, т. 1/1878, с. 135; год 11, т. 3/1878 с. xi), однако, полагал, что предпосылки следует изменять в соответствии с новыми данными, что следует всегда предпочитать простейшие законы, но только до тех пор, пока наблюдения не заставят отбросить их. И вот выдержка из последнего источника:

*Такова слабость человеческого разума, что он часто нуждается в помощи гипотез, чтобы соединить события друг с другом. Ограничивая этой целью применение гипотез и избегая придавать им реальность, которой они вовсе не имеют, и непрестанно исправляя их новыми наблюдениями, приходишь, наконец, к истинным причинам, или по крайней мере к законам явлений. История философии являет нам не один пример преимуществ, которые таким образом могут предоставить гипотезы.*

**3.1.3.** Лаплас изменял обозначения от одного сочинения к другому (а иногда и в пределах одного и того же труда) и их упорядочение оказалось слишком трудным; впрочем, в некоторых случаях нам пришлось ввести собственные обозначения. Далее, формула Лапласа вида $P(t) = f(t)$ означает, что для некоторой случайной величины $\xi$, которая может и не совпадать с $t$,
$P(t \leq \xi \leq t + dt) = f(t)dt$, где, по современным понятиям, второе нестрогое неравенство следовало бы заменить на строгое.

**3.2. Ранние мемуары.** Сочинения Лапласа по теории вероятностей легко разделить на две группы. В XVIII в. он исследовал пути применения сравнительно нового средства, плотности распределения, и сравнивал друг с другом различные более или менее естественные правила для выбора оценок истинных значений измеряемых постоянных[1]. Как и Даниил Бернулли (§ 2.8), он выводил исключительно сложные уравнения и вынужден был ограничиваться случаем трех наблюдений. Затем, однако, (нестрого) доказав несколько вариантов ЦПТ, Лаплас обратился к рассмотрению большого числа наблюдений. Послушаем Бьенеме (1853, с. 312):

*С 1770 по 1809 гг.* [с 1774 по 1811 гг.] ... *Лаплас давал многочисленные мемуары о вероятностях. Но, как бы они ни были интересны, он не хотел объединять их в общую теорию. Однако, как только он установил свойства функций вероятностей* [ЦПТ], *то ясно увидел, что это тот самый принцип, который управляет почти всеми приложениями, и составил свою теорию* [АТВ].

**3.2.1.** ***О вероятностях причин, определяемых по событиям*** (1774).

**3.2.1а.** Плотность распределения[2] ошибок наблюдения. Лаплас (с. 45) произвольно принимает для искомой кривой условие

$$\varphi'(x_2)/\varphi'(x_1) = \varphi(x_2)/\varphi(x_1)$$

и получает

$$\varphi'(x)/\varphi(x) = C,$$

$$\varphi(x) = (m/2)e^{-m|x|} \qquad (1)$$

или, мы бы сказали,

$$\varphi(x) = (m/2)e^{-m|x-h|}, \quad (1a)$$

где $h$ – параметр сдвига.

Но применяет Лаплас не эту кривую. Пусть $u$ и $v$ – две возможные оценки истинного значения постоянной и 3 (только 3) наблюдения, которые он молчаливо посчитал независимыми, расположены так, что **рисунок**

$$u - a = x,\ b - u = p - x,\ c - u = p + q - x,$$
$$v - a = x_1,\ b - v = p - x_1,\ c - v = p + q - x_1,\ p = b - a,\ q = c - b.$$

Тогда, как заявил Лаплас (с. 43), вероятности, что $u$ и $v$ действительно представляют искомую постоянную, относятся, в соответствии с принципом обращенной вероятности[3], как

$$\varphi(x)\ \varphi(p - x)\ \varphi(p + q - x) \div \varphi(x_1)\ \varphi(p - x_1)\ \varphi(p + q - x_1).$$

Он мог бы теперь ввести принцип наибольшего правдоподобия (§2.3), но поступил иначе.

**3.2.1b.** Оценка истинного значения измеряемой постоянной. Лаплас (с. 44) решил, что оценка ($e$) должна либо совпадать с медианой кривой[4]

$$y = f(x) = \varphi(x)\ \varphi(p - x)\ \varphi(p + q - x), - \quad (2)$$

где $x$ уже аргумент, а не некоторое расстояние, как в пункте А, – либо подчиняться условию

$$\int_{-\infty}^{\infty} |x - e|\, f(x)\, dx = \min, \text{ откуда } \int_{-\infty}^{\infty} f(x)\, dx = \int_{e}^{\infty} f(x)\, dx,$$

так что $e$ снова оказывается медианой. Мы можем назвать только двух авторов, применивших подобный выбор плотности, – Ньюкома (§ 6.5.3) и Питмена (1939), на которого сослался Эйзенхарт (1964).

Затем Лаплас вывел уравнение с неизвестным $x$, – расстоянием между искомой оценкой и $a$ (снова изменение обозначения!)

$$e^{mx} = e^{mp}[1 + (1/3)e^{-mp} - (1/3)e^{-mq}],$$

и при малых значениях $m$ оказалось, что $x \approx (2p + q)/3$ и оценка $e = x + a$ совпадала со средним арифметическим.

Лаплас (с. 48) отказывается от этого вывода, потому что, за исключением больших значений $|x|$, кривая (1) вырождается в прямую, что "вне всякого правдоподобия".

Его дальнейшие рассуждения не относятся к теории ошибок. Основным в них было то, что параметр $m$ неизвестен и наблюдения не оценивают его непосредственно. Случайная величина с плотностью

(1) имеет дисперсию $\sigma^2 = 2/m^2$, о которой Лаплас, конечно же, еще не знал.

**3.2.2.** ***О вероятностях*** (1781). Как и в 1774 г., Лаплас исходил из кривой типа (2). На этот раз он перечислил 4 возможных условия для выбора *среднего.*

$$\text{i)} \int_{-N}^{0} f(x)\,dx = \int_{0}^{N} f(x)\,dx \qquad (3)$$

где $N$— наибольшая возможная ошибка.

ii) То же уравнение с подынтегральной функцией $xf(x)$.

iii) Принцип наибольшего правдоподобия.

iv) Условие, "соответствующее сути задачи"

$$\int_{-N}^{N} f(x)\,dx = \min.$$

Оно, однако, совпадает с первым (3).

Далее Лаплас (§7) обсуждает специальную задачу, которая вначале, как представляется, не имеет отношения к теории ошибок, см. также его АТВ, § 15. Интервал а разделен на $i$ равных или неравных частей, из концов которых восставлены перпендикуляры, не возрастающие слева направо с суммой длин равной $s$. Их верхние концы образуют монотонно убывающую ломаную.

Если неоднократно повторять это построение, то среднее значение длины $r$-го перпендикуляра будет равно, как доказал Лаплас,

$$E\xi_r = (s/i)\,[(1/i) + 1/(i-1) + \ldots + 1/r].$$

Допустим, что эксперты располагают вероятности $i$ причин некоторого явления в убывающем порядке и что $s = 1$. Тогда $E\xi_r$ будет соответствовать вероятности $r$-й причины (или, при другом истолковании, – достоинств $r$-го кандидата на какую-либо должность). Но затем Лаплас предположил, что $i = a/dx$ и $r = x/dx$ и получил, уже в непрерывном случае,

$$(1/a)dx\int_{x}^{a} dx/x = (1/a)\ln(a/x),\ 0 < x \le a$$

и заметил, что $a$ является косвенной мерой точности наблюдений, а полученная логарифмическая кривая, или, точнее,

$$y = (1/2a)\ln(a/|x|),\ |x| \le a, \qquad (4)$$

представляет собой "средний закон ошибок"[5]. Его можно, разумеется, обобщить, включив параметр сдвига

$$y = (1/2a)\ln(a/|x-h|),\ |x-h| \le a. \qquad (4a)$$

О том, что функция (4) не существует при $x = 0$ (а функция (4а) – при $x = h$), Лаплас умолчал[6]. Он обосновал ее применение тем, что среднее следует принять, поскольку нет причин предпочитать что-то иное, ср. его выбор плотности в § 2.3.5 и наши общие замечания в § 3.1.2, и поскольку кривая (4) четная, определена только на конечном интервале и убывает с ростом $|x|$, т. е. соответствует свойствам "обычных" случайных ошибок.

Совершенно особо отметим, что, пытаясь определить некоторый параметр типа $h$ в (4а) из уравнения

$$\int_{-\infty}^{h} \varphi(x_1 - x)\, \varphi(x_2 - x) \ldots \varphi(x_n - x)\, dx =$$

$$\int_{h}^{\infty} \varphi(x_1 - x)\, \varphi(x_2 - x) \ldots \varphi(x_n - x)\, dx, \qquad (5)$$

где $x_1, x_2, \ldots, x_n$ обозначают результаты наблюдения, Лаплас (с. 480) ввел вместо (4а) последовательность функций

$$y = f(\beta x) = f(-\beta x) = \text{либо } q,\ \beta x = 0; \text{ либо } 0,\ \beta x \neq 0$$

при $\beta \to 0$ и доказал, что для $f[\beta(x - h)]$ это уравнение приводило к среднему арифметическому: $h = \bar{x}$. На *физическом* уровне он при этом ввел дельта-функцию Дирака, однако на языке обобщенных функций уравнение (5) не имеет смысла.

### 3.3. Дальнейшие работы 1810 – 1811 гг.

**3.3.1.** ***О приближении функций очень больших чисел ...*** (1810а). Фактически Лаплас рассматривал $n$ случайных ошибок (или величин) $\xi_i$, распределенных равномерно на интервале $[-h; h]$. Применив подобие характеристической функции и формулу обращения, он (с. 325, нестрого) доказал, что, при $n \to \infty$, в современных обозначениях,

$$\lim P(-s \le \bar{\xi} \le s) = (1/\sigma\sqrt{2\pi}) \int_{0}^{s} \exp(-x^2/2\sigma^2)\, dx, \qquad (6)$$

где $\sigma^2 = h^2/3$ – дисперсия каждой случайной величины.

Его рассуждения и формальные преобразования были исключительно небрежными, см. также Прим. 6, однако окончательная формула (6) верна. Он также обобщил свое изложение, рассмотрев произвольные, но одинаково распределенные величины, обладающие дисперсией.

**3.3.2.** ***Дополнение к предыдущему мемуару*** (1810b). Оно было посящено МНКв и написано сразу же после (и, видимо, ввиду) появления первого гауссова обоснования этого метода в 1809 г. Гаусс вывел нормальный закон, исходя из принципов среднего арифметического и наибольшего правдоподобия, Лаплас же предположил, что число наблюдений велико (этого Гаусс никогда не требовал) и что они разбиты на группы, средние из которых нормально распределены ввиду ЦПТ (см. пункт § 3.3.1). Он поэтому не ввел никакого предположения о среднем арифметическом, но его второе условие представляется искусственным.

**3.3.3.** ***Об определенных интегралах...*** (1811). Здесь Лаплас вернулся к МНКв, полагая, что он

> *до настоящего времени является полезным лишь в том, что позволяет без всякого гадания образовывать конечные* [нормальные] *уравнения ...,*

тогда как он, Лаплас, смог доказать, что МНКв "в то же время приводит к наиболее точным поправкам" (с. 362).

Несколько позже Лаплас (1814/1999, с. 862, правый столбец) по существу повторил это утверждение, но он мог бы быть намного более благожелателен по отношению к Гауссу, а его утверждение о наиболее точных поправках по меньшей мере не вполне определенно.

Лаплас (§ 6) вначале исследовал случай одного неизвестного

$$a_i + l_i = \varepsilon_i,\ i = 1, 2,\ldots, n. \quad (7)$$

Умножая эти уравнения, правые части которых – неизвестные ошибки свободных членов, на неопределенные множители $q_i$ и складывая произведения, он получает

$$[aq]x + [lq] = [eq],\ x = -[lq]/[aq] + [\varepsilon q]/[aq] = -[lq]/[aq] + s$$

и выводит распределение величины $s$, т. е. линейной формы ошибок $\varepsilon_i$. Молчаливо допустив, что $q_i$ имеют один и тот же порядок, он доказывает еще один вариант ЦПТ (ср. § 3.3.1), получив, опять-таки нестрого, в современных обозначениях

$$P(s = \alpha) = (1/\sigma_s\sqrt{2\pi})\ \exp(-\alpha^2/2\sigma_s^2),$$

$$\sigma_s^2 = k_2[qq]/[aq]^2,\ k_2 = \int_{-\infty}^{\infty} x^2\ \varphi(x)dx, \quad (8)$$

где $\varphi(x)$ — четная плотность распределения ошибок наблюдения. Далее Лаплас ввел "среднее значение ошибки, которой следует опасаться", т. е. ее абсолютное математическое ожидание, и потребовал (с. 393), чтобы оно была наименьшим. Это привело его к условиям

$$q_i = \mu a_i \quad (9)$$

и к принципу наименьших квадратов.

Абсолютное ожидание ошибки в качестве мерила он ввел намного раньше (§ 3.2.1 – 3.2.2), а позднее Гаусс (§ 5.3.2) заменил его дисперсией, тогда как Лаплас в основном держался своего собственного выбора.

В § 8 Лаплас рассмотрел случай косвенных наблюдений с двумя неизвестными

$$a_i x + b_i y + \alpha_i = \varepsilon_i,\ i = 1, 2, \ldots, s, \qquad (10)$$

где $\varepsilon_i$ снова были ошибками наблюдения. Последовательно умножив эти уравнения на неопределенные множители $m_i$ и $n_i$ и сложив полученные произведения, он получил

$$[am]x + [bm]y + [\alpha m] = [\varepsilon m],\ [an]x + [bn]y + [\alpha n] = [\varepsilon n].$$

Обозначим оценки $x$ и $y$ (которых Лаплас не ввел) через $\hat{x}$ и $\hat{y}$, тогда

$$[am]\Delta\hat{x} + [bm]\Delta\hat{y} = [\varepsilon m],\ [an]\Delta\hat{x} + [bn]\Delta\hat{y} + [ln] = [\varepsilon n].$$

Пусть величины $\varepsilon_i$ обладают четной плотностью и вторым моментом $k_2$. Тогда, как доказал Лаплас,

$$P([\varepsilon m] = l_1\ [\varepsilon n] = l_2) =$$

$$(1/4k_2\pi\sqrt{E})\exp\{-(1/4\,k_2E)\,[l_1^2[nn] - 2l_1l_2[mn] + l_2^2[mm]\}, \qquad (11)$$
$$E = [mm][nn] - [mn]^2.$$

Переходя от $l_1$ и $l_2$ к погрешностям $\Delta\hat{x}$ и $\Delta\hat{y}$ и интегрируя эти переменные в определенных пределах, Лаплас получает двумерное нормальное распределение и доказывает, что наименьшее абсолютное ожидание ошибок каждого неизвестного в отдельности приводит к МНКв.

Не вводя коэффициента корреляции, даже не упоминая, что $[\varepsilon m]$ и $[\varepsilon n]$ не являются независимыми, Лаплас сумел вывести верную формулу (11) используя аналог характеристической функции для двумерной случайной величины. Он не попал в ловушку определения предельных нормальных распределений для $[\varepsilon m]$ и $[\varepsilon n]$ и неверного установления их совместного распределения без учета указанной зависимости. Можно ли считать, что он всё это понимал, или же, что он просто не обращал внимания на упрощения или достижение единообразия? В соответствии с § 3.8.1 мы полагаем, что второе было также возможным.

Позднее Лаплас (§ 3.7.2), основываясь на формуле (11), по существу заявил, что сумма двух нормальных распределений также нормальна. Это утверждение было известно и Гауссу (§ 5.1.5), но доказал его Бессель (§ 6.1.1d).

**3.4.** ***Аналитическая теория вероятностей*** (1812). Теория ошибок представлена там главой 4-й, хорошо известной своей трудностью.

**3.4.1.** В §§ 18,19 и 22 Лаплас изучал распределения различных функций ошибок наблюдения $\varepsilon_i$, одинаково распределенных на конечном интервале, – $\sum\varepsilon_i$, $\sum|\varepsilon_i|$, $[\varepsilon\varepsilon]$ и $[q\varepsilon]$, где $q_i$ были числа одного и того же порядка. Он уже раньше рассматривал сумму большого числа ошибок (§ 3.3.1) и здесь, в АТВ, он применил тот же самый подход во всех случаях.2

В § 19 Лаплас получил

$$P(-\,n^2r\sqrt{s} \le [\varepsilon\varepsilon] - 2\,k_2n^2s/k \le n^2r\sqrt{s}\,) = (\beta/\sqrt{\pi})\int_0^r \exp(-\,\beta^2x^2/4)dx \qquad (12)$$

где $\beta = k/\sqrt{kk_4 - k_2^{\,2}}$.

Он при этом ввел обозначения

$$k = \int_0^1 \varphi(x)\,dx,\; k_2 = \int_0^1 x^2\varphi(x)\,dx,\; k_4 = \int_0^1 x^4\varphi(x)\,dx$$

и предположил, что погрешности $\varepsilon_i$, $i = 1, 2, ..., s,$ обладают четной плотностью $\varphi(x/n)$ при $|x| \le n$. Заметим, что

$$\mathrm{E}\varepsilon_i^2 = 2n^2\,k_2/k,\; \mathrm{E}\varepsilon_i^4 = 2\,n^5\,k_4,\; \mathrm{D}\varepsilon_i^2 = 2n^4/\beta^2.$$

В § 22 Лаплас вывел нормальный закон N (0; σ) для линейной формы $\xi = [q\varepsilon]$. Одну и ту же плотность ошибок он предположил четной при $\mathrm{E}\varepsilon_i^2 = k_2$. Тогда, в современных обозначениях, оказалось, что $\sigma^2 = k_2[qq]$.

**3.4.2.** В §§ 20 — 21 Лаплас повторил свое прежнее исследование (§ 3.3.3), исходя из уравнений (7) и (10) и снова предположил большое число наблюдений.

**3.4.3.** В § 23 Лаплас (с. 338) попытался исследовать "средний результат, который многочисленные и еще не сделанные наблюдения должны указывать с большей пользой ...". Та же идея привела Симпсона (§ 2.6) к введению кривых плотности, но Лаплас высказал ее явно. Впрочем, ничего нового он уже не достиг и признал, в том же параграфе, что ничего кроме МНКв предложить не может[7].

Пусть истинное значение измеряемой постоянной равно $A + w$ ($w > 0$), а результаты ее $s$ наблюдений равны $A,\ A + q_1,\ A + q_2,...,$ $A + q_{s-1}$ ($q_1 \le q_2 \le \ldots \le q_{s-1}$) и $\varphi(x) = \exp\,[-\,\psi(x^2)]$ – плотность их погрешностей. Снова (§ 3.2.1) применяя принцип обращенной вероятности, Лаплас замечает, что

$$P(w) = y = H\varphi(-w)\,\varphi(q_1 - w)\,\varphi(q_2 - w)\ldots\varphi(q_{s-1} - w). \qquad (13)$$

Обозначим оценку $w$ через $e$. Тогда, как решил Лаплас (ср. § 3.2.2),

$$\int_0^e (e - x)ydx = \int_e^\alpha (x - e)ydx, \int_0^e ydx = \int_e^\alpha ydx, \qquad (14a;\ 14b)$$

где α – "конец" кривой (13), т. е. число, вряд ли превышающее $q_{s-1}$.

Забыв про условие (9), Лаплас (с. 340) утверждает, что уравнение (14а) "очевидно указано теорией вероятностей" и потому предпочтительнее, чем принцип наибольшего правдоподобия, который рекомендовали "прославленные ученые" [Ламберт, §2.4.3, Даниил Бернулли (§ 2.8.2) и Гаусс в 1809 г. (§ 5.1.2)][8]. Впрочем, для выбранной им четной [и одновершинной] плотности никакого отличия между указанными принципами не было. Лаплас применил второй, но добавил гауссов постулат арифметического среднего (§5.1.2).

Пусть $w = a + z$, где $z$ – небольшая поправка. Тогда выражение (13), которое принимает вид

$$y = H\exp(-M - 2Nz - Pz^2 - \ldots),$$

окажется максимальным при $N = 0$, т. е. если

$$a\psi'(a^2) + (a - q_1)\,\psi'[(a - q_1)^2] + \ldots + (a - q_{s-1})\,\psi'[(a - q_{s-1})^2] = 0.$$

Пусть теперь результаты $i$ наблюдений равны $A$, а результаты остальных – $A + q$, тогда среднее арифметическое, "принятое наблюдателями" (с. 342), будет равно $a = (s - i)q/s$ и

$$\psi'(a^2) + (s - i)\,(a - q)\,\psi'[(a - q)^2] = 0,$$

$$\psi'[(s - i)^2q^2/s^2] = \psi'(i^2q^2/s^2) = g,$$

$$\varphi(x) = \exp(L - gx^2) = \sqrt{g/\pi}\,\exp(-gx^2).$$

Замечая, что при этой плотности возможны ошибки любой величины, Лаплас (с. 343) добавляет странное утверждение: это свойство не является существенным, потому что "можно принять $g$ достаточно большим", так что вне пределов допустимых ошибок $\varphi(x)$ окажется пренебрегаемой[9]. К тому времени Гаусс (1809b, §178) отметил, что при любых практически интересных значениях $g$ крупные погрешности возможны лишь с пренебрегаемыми вероятностями и что $\sqrt{g}$ – мера точности наблюдений (а не произвольно выбираемая величина).

**3.4.4.** Последний параграф (§ 24) посвящен методу минимакса (ср. наш § 1.9) и истории МНКв. И здесь Лаплас (с. 353) признал фактический (но не формальный) приоритет Гаусса:

> *Г-н Лежандр возымел простую идею рассматривать сумму квадратов ошибок наблюдений и приводить ее к минимуму, что непосредственно приводит к стольким же окончательным* [нормальным] *уравнениям, сколько элементов следует исправить. Этот ученый геометр первым опубликовал указанный метод, но следует отдать должное г-ну Гауссу, заметив, что за много лет до публикации Лежандра он постоянно пользовался той же идеей и сообщил о ней многим астрономам.*

### 3.5. *Дополнение 1* (1816) *к Аналитической теории вероятностей*

**3.5.1.** Рассмотрим исходные уравнения с двумя неизвестными

$$a_ix + b_iy + l_i = v_i,\ i = 1, 2, ..., s. \tag{15}$$

Истинные ошибки наблюдений будут равны

$$\varepsilon_i = v_i - (a_i\Delta x + b_i\Delta y) \qquad (16)$$

где $\Delta x$ и $\Delta y$ – погрешности $x$ и $y$, определяемые по МНКв (в ином обозначении – $\Delta\hat{x}$ и $\Delta\hat{y}$). Тогда

$$[\varepsilon\varepsilon] = [vv] + \sum_{i=1}^{s} (a_i\,\Delta x + b_i\Delta y)^2. \qquad (17)$$

Введем вслед за Лапласом

$$\Delta x = \xi/\sqrt{s},\ \Delta y = \eta/\sqrt{s}, \qquad (18)$$

$$t = [vv]/\sqrt{s} - 2k_2n^2\sqrt{s}/k, \qquad (19)$$

где обозначения те же, что в формуле (12)[10], и пусть

$$Q^2 = \sum_{i=1}^{s} (a_i\xi + b_i\eta)^2. \qquad (20)$$

Локальная теорема, соответствующая интегральному выражению (12) приводит к

$$P([\varepsilon\varepsilon] - 2k_2n^2s/k = n^2r\sqrt{s}) \approx \exp(-\beta^2r^2/4),$$

или, с учетом уравнения (17) – (20), левую часть можно записать в виде

$$P\{r = (1/n^2)\,[t + (Q^2/s\sqrt{s})]^2\}.$$

В обозначениях Лапласа

$$P(t) \approx \exp\{-(\beta^2/4n^4)\,[t + (Q^2/s\sqrt{s})]^2\} \qquad (21)$$

и кроме того

$$P(\xi;\eta) \approx \exp\{-Q^2/(2[vv] - 2t\sqrt{s})\}, \qquad (22)$$

что может быть получено из формул (17) – (20) и выражения

$$P(\Delta x;\Delta y) \approx \exp\Big[-\frac{k}{4k_2n^2}\sum_{i=1}^{s} (a_i\Delta x + b_i\Delta y)^2\Big]. \qquad (23)$$

Лаплас фактически не доказал выражения (23); это сделал Тодхантер (1869) и, более простым способом, Медоукрофт (1920).

Лаплас далее утверждает, что $P(\xi;\ \eta;\ t)$ пропорционально произведению правых частей выражений (22) и (21) и замечает, что эту вероятность можно применить для оценки $\xi$ (или $\eta$). Если умножить $P(\xi;\ \eta;\ t)$ на $d\xi\ d\eta\ dt$ и проинтегрировать произведение, во-первых по всем возможным значениям $\eta$ и $t$ и по некоторому интервалу $[\alpha;\ \beta]$ величины $\xi$; и, во-вторых, по всем возможным значениям всех трех переменных и разделить первый интеграл на второй, то частное окажется равным вероятности $P(\alpha \le \xi \le \beta)$.

Формула Лапласа для $P(\xi; \eta; t)$ показывает, что он молчаливо принял, что при большом числе независимых наблюдений с четной плотностью, которая в пределе становится нормальной и потому подчиняется формуле (12), $(\xi; \eta)$ не зависит от $t$. Он, стало быть, основывал свои выводы на недоказанной им теореме, которая представляет собой его форму утверждения о независимости выборочной дисперсии и среднего для независимых и нормально распределенных ошибок наблюдения, ср. § 5.3.4. Лаплас, однако, не подчеркивал своего способа оценки $\xi$ или $\eta$, он скорее рекомендовал попутно, при решении соответствующей системы нормальных уравнений, определять вес оценки. Это действительно возможно, но только для того неизвестного, которое остается после исключения из системы всех остальных. Если же угодно определить вес другой оценки, например $\eta$, то, продолжал Лаплас, следует снова решать систему нормальных уравнений при условии, что $y$ будет последним неизвестным в редуцированной системе. Мы опишем отрицательное мнение Гаусса об этом способе в Прим. 9 к гл. 5.

**3.5.2.** Лаплас обобщает свое изложение на случай не четных плотностей, для которых $\sum \varepsilon_i \neq 0$. В соответствии с уравнением (16)

$$\sum_{i=1}^{s} \varepsilon_i = \sum_{i=1}^{s} [v_i - (a_i\Delta x + b_i\Delta y)].$$

Исследовав соответствующие приближения, он утверждает, что в этом случае можно потребовать, чтобы

$$\sum_{i=1}^{s} v_i = 0. \qquad (24)$$

При $a_i$ = Const это условие выполняется первым нормальным уравнением $[av] = 0$; при $b_i$ = Const, – вторым, $[bv] = 0$ и т. д., см. § 4.4. В противном случае, однако, (24) можно считать дополнительным нормальным уравнением, соответствующим фиктивному, но существенному неизвестному (среднему значению систематической ошибки, присутствие которой привело к нарушению четности плотности), коэффициент при котором во всех исходных уравнениях (15) был одним и тем же.

### 3.6. *Дополнение 2* (1818) *к Аналитической теории вероятностей*

**3.6.1.** Лаплас (с. 536) здесь признает нормальное распределение не только как предельное (что он делал с 1810 г.), но и как действительный закон ошибок наблюдений:

> *Это предположение, самое естественное и самое простое из всех, следует из применения повторительных теодолитов при измерении углов в триангуляции.*

Действительно (§ 9.2), повторительный теодолит позволял уменьшить полную ошибку измерения угла, притом в большой степени уравнивал влияние двух основных ее источников. Однако, другие (очевидно менее существенные) ошибки оставались прежними, и условия для приложимости ЦПТ, и, стало быть, для появления нормального распределения полной ошибки наблюдения вряд ли выполнялись.

**3.6.2.** Лаплас рассмотрел уравнивание наблюденных углов треугольника. Пусть их ошибки равны α, β и γ, а их нормальное распределение

$$\varphi(x) = \sqrt{h/\pi}\, \exp\,(-\,hx^2).$$

Тогда, если $T$ – невязка треугольника,

$$P(\alpha;\, \beta;\, \gamma) \approx$$
$$\exp\,\{-2h[\beta + (1/2)\alpha - (1/2)T]^2 - (3h/2)[\,\alpha - (1/3)T]^2 - (h/3)T^2\}.$$

Интеграл от $P(\alpha;\, \beta;\, \gamma)d\beta$ по интервалу $-\infty < \beta < \infty$ имеет множителем

$$\exp\,[-\,(3h/2)[\alpha - (1/3)T]^2 - (h/3)T]^2,$$

так что вероятнейшее значение α равно $T/3$. Соответственно, Лаплас решает, что каждый угол следует исправлять на одну и ту же величину, $T/3$.

МНКв приводит к тому же выводу, притом для всякой плотности ошибок наблюдений. Совет Лапласа не имеет поэтому никакого значения, притом что условия при уравнивании цепи триангуляции включают и те, которые соответствуют наличию двух базисов (и, возможно, двух астрономических азимутов) на обеих ее концах[11] и предварительное уравнивание треугольников вовсе не обязательно.

Обозначим $\alpha - T/3 = \alpha_1$. Тогда, по Лапласу,

$$P(\alpha_1; T) \approx \exp[-\,(h/3)T^2 - (3h/2)\alpha_1{}^2]$$

и для цепи из $n$ треугольников

$$P(T_1,\, T_2,\, \ldots,\, T_n) \approx (h/3\pi)^{n/2} \exp\,\{-\,(h/3)\,[TT]\},$$

$$P(h) = h^{n/2} \exp\,\{-\,(h/3)\,[TT]\} \div \int_0^\infty h^{n/2} \exp\,\{-\,(h/3)\,[TT]\}dh, \quad (24a)$$

$$Eh = \int_0^\infty hP(h)\, dh = \{(3n + 2)/2[TT]\} \approx 3n/(2[TT]). \qquad (24b)$$

Заметим, что при $h = Eh$

$\sigma = 1/\sqrt{2h} = \sqrt{[TT]/3n}$

и что Лаплас рассматривал $h$ как случайную величину.

**3.6.3.** Попытка улучшить метод наименьших квадратов (с. 563). Пусть исходные уравнения с одним неизвестным имеют вид

$$p_i y - l_i = v_i,\ i = 1, 2, \dots, n. \tag{25}$$

По МНКв

$$\hat{y}_1 = [lp]/[pp], \tag{26}$$
$$\Delta\hat{y}_1 = [pv]/[pp] \equiv u_1,$$

но можно также предположить, что $\sum v_i = 0$ и тогда, при умножении уравнений (25) на неопределенные множители $m_i$ и сложении произведений,

$$\hat{y}_2 = [ml]/[mp],\ \Delta\hat{y}_2 = [mv]/[mp] \equiv u_2.$$

Для нормально распределенных ошибок совместное распределение $\Delta\hat{y}_1$ и $\Delta\hat{y}_2$ пропорционально

$$\exp\{-(k/2k_2E)([pv]^2[mm] - 2[pv][mv][mp] + [mv][pp])\} =$$
$$\exp\{-k[pp]/(2k_2E)[u_1^2E + (u_2-u_1)^2[mp]^2]\}, \tag{27}$$
$$E = [mm][pp] - [mp]^2.$$

Далее, $(u_2 - u_1) = (\hat{y}_1 - \hat{y}_2)$, так что эта разность известна из наблюдений и искомая вероятность может изменяться только в зависимости от $u_1$. Она становится наибольшей, когда эта величина равна нулю и, другими словами, никакая поправка к оценке по МНКв не уменьшает ее дисперсии.

Формула Лапласа (27) основана на его предыдущей работе (§3.3.2), в которой он смог вывести совместное распределение двух линейных форм тех же самых ошибок. Он не указал и не принял во внимание, что плотность ошибок $\varepsilon_i$, см. формулу (10), не совпадает с распределением остаточных свободных членов $v_i$ в уравнениях (25). По крайней мере при решении этих уравнений по МНКв, см. формулу (26), это отличие не должно быть забыто. Чубер (1890) доказал, что при уравнивании непосредственных наблюдений по МНКв в предположении нормального распределения ошибок $\varepsilon_i$ имеющих плотность

$$\varphi(x) = (h_1/\sqrt{\pi})\exp(-h_1^2x^2)$$

остаточные свободные члены имеют нормальное распределение с мерой точности $h_2 = h_1\sqrt{n/(n-1)}$. Случай ошибок $\varepsilon_i$, не подчиняющихся нормальному закону, очевидно более сложен.

**3.6.4.** Метод Бошковича уравнивания наблюдений (см. § 2.5.1).

При решении системы уравнений

$$a_i x + y + l_i = v_i,\ i = 1, 2, ..., n$$

Бошкович принял два условия

$$v_i + v_2 + ... + v_n = 0,\ |v_1| + |v_2| + ... + |v_n| = \min. \qquad (28a; 28b)$$

Первое из них приводит к

$$(a_i - \bar{a})x + (l_i - \bar{l}) = v_i, \qquad (29)$$

откуда, ввиду (28b), следует медиана из всех подобных отношений, т. е. из $-(l_i - \bar{a})/(a_i - \bar{l})$.

Лаплас (1792, с. 506 – 516; *Неб. Мех.*, т. 2, 1798, § 40) применил этот "остроумный" метод (1792, с. 506), а его переводчик, Боудитч, сравнил методы Бошковича и наименьших квадратов (§ 2.5.1). Лаплас (1818, с. 571 – 580) также теоретически исследовал метод Бошковича, называя его "методом ситуации" (с. 576), но уже не упоминая Бошковича. Ниже и в § 3.6.5 мы следуем его изложению. Выпишем уравнения (29) с неизвестным $y$ в его форме:

$$p_i y - a_i = v_i,\ i = 1, 2, ..., n,\ p_i > 0.$$

Тогда

$$\hat{y} = a_j/p_j,\ \Delta\hat{y} = \varepsilon_j/p_j.$$

Если $a_1/p_1 > a_2/p_2 > \ldots > a_n/p_n$, то

$$p_1 + p_2 + \ldots + p_{j-1} < p_j + p_{j+1} + \ldots + p_n,$$
$$p_1 + p_2 + \ldots + p_j > p_{j+1} + p_{j+2} + \ldots + p_n. \qquad (30)$$

Действительно, если видоизменить доказательство Лапласа, можно заметить, что условия (30) приводят к минимуму функцию

$$W = p_1|y - (a_1/p_1)| + p_2|y - (a_2/p_2)| + \ldots + p_n|y - (a_n/p_n)|.$$

.

Пусть $v_i$ имеют четную плотность $\varphi(x)$, тогда для $a_k = p_k v_j/p_j$ и $v_j > 0$

$$P(v_s/p_s < v_j/p_j < v_t/p_t,\ s = 1, 2, ..., j-1,\ t = j+1, j+2, ..., n) \approx$$
$$\prod_{k=1}^{j-1}\left[1 - 2\int_0^{\alpha_k}\varphi(x)dx\right]\prod_{k=j+1}^{n}\left[1 + 2\int_0^{\alpha_k}\varphi(x)dx\right],$$

притом

$$P(v_j) = 1 + v_j^2\,\varphi''(0)/2\varphi(0).$$

Логарифм произведения этих двух вероятностей будет пропорционален

$$-2\beta\varphi(0)[p_1 + p_2 + \ldots + p_{j-1}) - (p_{j+1} + p_{j+2} + \ldots + p_n)] -$$

$$2\beta^2\varphi^2(0)[pp] + (1/2)p_j^2\{[\varphi''(0)/\varphi(0)] + 4\varphi^2(0)\}, \qquad (31)$$

где $\beta = v_j/p_j$. При большом $n$ с учетом неравенств (30) Лаплас получает взамен (31)

$$\exp\{-2\beta^2[pp]\varphi^2(0)\},$$

так что, добавим мы, $\sigma_{\hat{y}}^2 = 1/\{4[pp]\varphi^2(0)\}$, и сравнивает эту вероятность с соответствующей вероятностью для МНКв, при котором

$$\hat{y} = [ap]/[pp]$$

и, ср. формулу (8),

$$\sigma_{\hat{y}}^2 = k_2/[pp],\ k_2 = \int_{-\infty}^{\infty} x^2\varphi(x)\,dx.$$

Отсюда следует, что метод ситуаций предпочтительнее МНКв (а медиана предпочтительнее среднего арифметического), если

$$4\varphi^2(0) > 1/k_2.$$

Лаплас не отличал плотностей остаточных свободных членов $v_i$ (которые он вначале назвал "уклонениями", а затем "ошибками", см. его с. 571 – 572) и истинных ошибок $\varepsilon_i$, ср. наше соответствующее замечание в конце § 3.6.3.

Колмогоров (1931) заново сравнил обе указанные оценки. Обозначим выборочную и теоретическую медианы через $m_n$ и $m$ соответственно, пусть $n$ будет числом наблюдений, $\varphi(x)$ –непрерывной одновершинной плотностью распределения ошибок наблюдений, $\varphi(m) > 0$, $\sigma^2$ – вторым моментом и $\mu_n = (m_n - m)/\sqrt{n}$.

Тогда, как он доказал, при $n \to \infty$ распределение $\mu_n$ стремится к $N(0; \sigma_m)$, $\sigma_m = (1/2)\varphi(m)$. Таким образом, $\sigma_m$ можно сравнить с $\sigma$. Случай конечного $n$ также может быть изучен, поскольку медиана является одной из порядковых статистик[12].

**3.6.5.** Сочетание двух оценок. Как и в § 3.6.3, Лаплас рассматривает возможность сочетания двух оценок, на этот раз среднего арифметического и медианы, в случае большого числа наблюдений. Он определял такое значение $c$, для которого дисперсия величины $y_{MLSq} - c\,(y_{MLSq} - y_{MS})$ окажется минимальной (MS – метод ситуаций):

$$\sigma^2 = (1 - c^2)\sigma_1^2 + c^2\sigma_2^2 + 2c\,(1 - c)\,\mathrm{cov}\,(\varepsilon_1\,\varepsilon_2) = \min.$$

Здесь $\sigma_1^2$ и $\sigma_2^2$ – соответствующие дисперсии, а $\varepsilon_1$ и $\varepsilon_2$ – ошибки оценок. Лаплас действительно установил совместное распределение этих ошибок, что было необходимо для подсчета ковариации (современные понятие и термин) и определил $c$ (Хальд 1998, § 20.12). Для нормального распределения величин $v_i$ (более правильно: ошибок наблюдений) оказалось, что $c = 0$, так что никаких поправок не требовалось.

**3.7.** ***Дополнение 3*** **(примерно 1819) к** ***Аналитической теории вероятностей***

**3.7.1.** Лаплас оценивал точность цепи триангуляции из 26 треугольников, включенной в общую цепь, состоящую из 107 треугольников. Как и в § 3.6.1, он полагал, что ошибки наблюдения следуют нормальному закону

$$\varphi(x) = \sqrt{k/\pi}\,\exp(-kx^2), \tag{32}$$

так что (см. ниже)

$$\varepsilon = \int_{-\infty}^{\infty} |x|\varphi(x)dx = 1/\sqrt{\pi k}\,,\ \varepsilon' = \int_{-\infty}^{\infty} x^2\,\varphi(x)\,dx = 1/2k.$$

Пусть невязка треугольника $i$ равна $T_i$ (секунд дуги) и, стало быть, $[TT]$ характеризует точность цепи. Впрочем, Лаплас (ср. § 5.4.2) полагал, что более надежной оценкой точности окажется 28/107 той же оценки для всей цепи (что могло быть ошибочным ввиду различий местных атмосферных условий). Итак, он принял

$$[TT]_1 = (26/107)\,[TT]_2 = 108.184.$$

Затем Лаплас вычислил выборочное значение $\varepsilon$,

$$(1/107)\,(|T|_1 + |T|_2 + \ldots + |T|_{107}) = 1.62$$

и дисперсию (для случая нормального распределения) $\varepsilon' = \pi\varepsilon^2/2 = 4.13$ и заново получил $[TT]_1 = 4.13 \cdot 26 = 107.78$.

Близость двух значений $[TT]$ (т. е. наличие нормального распределения) Лаплас посчитал примечательным, но никаких количественных критериев (не известных в то время) не предложил. Он (с. 585) также упомянул дисперсию:

> *Можно оценить относительную точность инструментов, применяемых в геодезических наблюдениях, по величине* $\varepsilon'$*, выведенной по большому числу треугольников.*

По существу Лаплас оценивал точность дисперсией еще раньше (§§3.6.4 и 3.6.5), но лишь в случае нормальных распределений, здесь же он никаких оговорок не сделал.

**3.7.2.** Наблюдения, искаженные несколькими источниками ошибок. Окончание *Дополнения* посвящено решению системы уравнений с одним неизвестным ($y$):

$$p_i y = a_i + m_i \gamma_i + n_i \lambda_i, i = 1, 2, \ldots, n. \quad (33)$$

Ошибки $\gamma_i$ и $\lambda_i$ полагаются нормально распределенными с какими-то параметрами ($k_{11}$; $k_{21}$) и ($k_{12}$; $k_{22}$):

$$\varphi(x) = (k/\sqrt{\pi}) \exp(-k^2x^2). \quad (34)$$

Заметим, что эта функция не совпадает с (32). Легко видеть (см. ниже), что $k_{11} = k_{12} = 1$.

Умножив каждое уравнение (31) на $f_i$ и сложив произведения, он получил

$$\hat{y} = [af]/[pf] + ([mf\gamma] + [nf\gamma])/[pf],$$

где второе слагаемое в правой части является ошибкой $\Delta\hat{y}$, которую он обозначил через $s$. Далее Лаплас (с. 609) заявил, что "по предыдущему параграфу"

$$P(s) \sim \exp\{\frac{-s^2[pf]^2}{4k_{12}[mf]^2 + 4k_{22}[nf]^2}\}.$$

Хотя предыдущий параграф (с. 601 – 603) не поясняет этого, но очевидно, что Лаплас представлял себе, что сумма нормальных законов снова нормальна, ср. §§ 5.1.5 и 6.1.1d.

Наконец, он (с. 609) выбрал такие значения $f_i$, при которых $P(s)$ минимальна при заданном $s$ (точнее: при которых минимальна дисперсия $s$):

$$f_i = p_i/(k_{12} m_i^2 + k_{22} n_i^2). \quad (35)$$

Это исследование Лаплас обобщил на случай двух неизвестных; позднее он (1827, с. 349) заметил, что ошибки $\gamma_i$ и $\lambda_i$ предполагались независимыми.

Также в 1827 г. Лаплас (с. 346) применил тот же прием в случае ошибок с различными четными, но не обязательно нормальными распределениями[13] и без доказательства указал, что формула (35) по-прежнему действительна. Да, эта формула соответствует вычислению соответствующих дисперсий. И вот комментарий Лапласа (с. 343):

> *Я определил с особой тщательностью те множители, на которые следует умножить различные исходные уравнения, чтобы вывести наиболее благоприятные результаты, при которых средняя ошибка, которой следует опасаться в избытке и недостатке, минимальна. Эти множители вовсе не те, к которым приводит прием, известный под названием метод наименьших квадратов, который является лишь частным случаем наиболее благоприятного метода и от которого они отличаются в большинстве случаев, когда их применяют. ...*

*Метод наименьших квадратов, который многие* [?] *геометры обосновали весьма малоудовлетворительно, вовсе не приводит здесь к наиболее выгодным множителям; он имел лишь то преимущество,* [Лаплас по сути повторяет свое раннее высказывание, см. § 3.3.3]. *Я представил* [следует ссылка на *Дополнение 3*] *общее выражение для наиболее благоприятных множителей*.

Мы вернемся к обоснованию МНКв в §§ 5.5.1 и 5.3, но некоторые замечания сформулируем здесь.

A) Задача Лапласа относилась к ошибкам, обладающим определенными плотностями, и в этом смысле не была общей.

B) При использовании МНКв уравнения (33) должны быть взвешены (если, конечно, это возможно), т. е. умножены на множители (35). Лаплас показал это до публикации Гаусса (1823b) с его вторым обоснованием МНКв. К 1827 г. формула (35), однако, была известна, притом в более широком смысле (не обязательно для нормальных распределений).

C) До 1823 г. цитированное выше мнение Лапласа было частично верно, но в 1827 г. уже нет.

D) Так же, как и раньше (§ 3.3.3), Лаплас мог бы быть намного более благожелателен по отношению к Гауссу.

**3.8. Ошибочное исследование зависимых наблюдений** (1827)**.** В § 3.3 мы описали как Лаплас успешно справился со случаем зависимых наблюдений, теперь же мы отметим его ошибки в аналогичном случае при исследовании влияния Луны на атмосферные приливы[14].

**3.8.1.** Сезонные и суточные вариации атмосферного давления. По наблюдениям, длившимся 11 лет (132 месяца), его среднесуточная вариация в Париже составила 0.763*мм*, а только лишь за периоды с февраля по апрель в те же годы – 0.940*мм*[15]. Было ли расхождение значимо? Лаплас предположил, что вероятности ошибок были $u$ и $u_1 = u + z$,

$$P(u) \sim \exp(-132\alpha^2u^2),\ P(u+z) \sim \exp[-33\,\alpha^2(u+z)^2],$$

где α определяется из самих наблюдений.

Далее,

$$P(u, u+z) = P(u)\cdot P(u+z), \qquad (36)$$

$$P(z) = \int_{-\infty}^{\infty} P(u, u+z)\,du$$

и поэтому вероятность $P(z \geq 0.940 - 0.763 = 0.177)$ может быть вычислена.

Неверной была формула (36): величины $u$ и $(u + z)$ не независимы. Раннее Лаплас (1814/1999, с. 836, левый столбец) указал, что независимость событий необходима для (обычной) теоремы умножения вероятностей, а позднее (1818, с. 534 – 535 и 561) упомянул независимость в связи с конкретными геодезическими задачами.

**3.8.2.** Влияние Луны. Лаплас также попытался определить момент $\lambda$ суток для максимального атмосферного прилива и влияние Луны $R$ на него. Введя два других неизвестных

$$x = 4R\sin 2\lambda,\ y = 4R\cos 2\lambda,$$

он составил уравнения

$$x\cos(2iq) + y\sin(2iq) = E_i,\ y\cos(2iq) - x\sin(2iq) = F_i,\ i = -1, 0, 1, 2,$$

в которых буквы $i$ указывали дни наблюдения, а $q$ было известной величиной. Лаплас решил эти уравнения, полагая, что свободные члены $E_i$ и $F_i$ независимы, что было явно неверно.

**Примечания**

**1.** Он (1781, §7; 1812, §15) также повторил вычисления Симпсона (§ 2.6) и Лагранжа (§ 2.7), относящиеся к непрерывным равномерному и треугольному распределениям.

**2.** Лаплас не ввел здесь никаких особых терминов. Позже он применял несколько различных словообразований и, наконец (1812; 1816; 1818), остановился на *законе вероятностей* (и *законе ошибок*). Напомним (§ 2.7) термин Лагранжа *кривая возможностей ошибок*.

**3.** Фактически принцип Бейеса с равномерным априорным распределением. Позднее Лаплас (1814/1999, с. 836, правый столбец) включил его в основные принципы теории вероятностей. Его применению Лаплас посвятил мемуар (1774), но до 1814 г. (1814/1999, с. 862, левый столбец) Бейеса он не упоминал.

**4.** Этот термин ввел Курно (1843, § 68). Лаплас (§ 3.6.4) впоследствии вновь исследовал применение медианы.

**5.** Каждую кривую, которая соответствует некоторой ломаной, можно понимать как реализацию случайного процесса, а эту среднюю кривую – как его ожидание.

**6.** На такие *мелочи* Лаплас не обращал внимания.

**7.** Следуя за Чебышевым (посм. публ. 1936, с. 227), Марков ввел возможные наблюдения уже в своих литографически изданных лекциях, а затем и в руководстве (1900; с. 323 и 373 в изд. 1924 г.), не предполагая никакой определенной плотности. Его бывший студент, Коялович, впоследствии автор своего собственного литографированного курса теории вероятностей, признался, что никогда этого не понимал, но что положение было бы иным, будь принято существование плотности. Это свидетельствует о том, что соотношение между случайными ошибками и случайными величинами следовало подчеркивать. Два письма Кояловича Маркову хранятся в Архиве РАН (фонд 173, оп. 1, № 10). Вот выдержка из одного из них (1893):

*Насколько я Вас понял, Вы рассматриваете каждое отдельное наблюдение как одно из значений возможного результата. Таким* *образом, для каждого измерения возможен ряд результатов …, один из которых осуществляется на деле. Всё это я готов понять для одного измерения, но когда их имеется, напр., два, то я не могу понять, чем отличается ряд возможных результатов первого наблюдения от ряда возможных результатов … второго измерения. Вопрос, конечно, сейчас же решается, если Вы скажете, что вероятность одной и той же ошибки в этих рядах различна, но ведь Вы, вероятно, не захотите вводить понятие о вероятности ошибки в Ваше изложение.*

**8.** Раннее Лаплас (1812, с. 352) назвал этих ученых по имени. Эйлер (§ 2.9), однако, не соглашался с принципом наибольшего правдоподобия, а Гаусс (§ 5.3.2 и прим. 12) впоследствии заявил, что он недостаточно хорош.

**9.** Раннее Лаплас (1774, с. 46) сам заявил, что (даже) случае распределения (1) вероятность крупных ошибок ничтожна.

**10.** В другом месте Лаплас (1818, с. 570) указал, что при $\Delta x$ и $\Delta y$, превышающих $1/\sqrt{s}$, результаты окажутся неудовлетворительными. Пусть уравнения (15) содержат одно неизвестное и притом $a_i = a =$ Const. Тогда $\Delta x = \sum \varepsilon_i / as \approx (\varepsilon_i / a)\sqrt{s}$ и можно будет надеяться, что $|\Delta x| < \sqrt{s}$.

**11.** Лаплас ни разу не вводил этих условий; уравнительные вычисления по существу начали Гаусс, его студенты (Герлинг) и Бессель.

**12.** Этьен (1926 – 1927) пылко защищал медиану, противопоставляя ее среднему арифметическому. Его сочинение содержало странные высказывания; он утверждал, например, что какие-то неназванные “отжившие догмы” еще живы лишь благодаря авторитету Гаусса (с. 422, прим.) и что МНКв *сегодня* столь же произволен, как и во времена Лежандра (с. 440, прим.). Он сам (с. 427) полагал, что у *добросовестного наблюдателя* каждое измерение может с одной и той же вероятностью быть ошибочным в каждую сторону. Намного более ранняя работа Этьена (1890) столь же отклонялась от общепринятых положений. На нее сослался Менделеев (1875, с. 209), см. § 6.3.1. Он же (1895, с. 159) ошибочно утверждал, что

*Из разнообразных определений можно, а иногда и должно, брать среднее только тогда, когда относительное достоинство определений или совершенно неизвестно или ничем ясно не определяется.*

Вот в таких случаях как раз и нужна медиана, мало зависящая от нескольких недоброкачественных измерений.

**13.** Более точно, на этот раз Лаплас принял, что $p_i =$ Const.

**14.** Также см. Стиглер (1986, с. 148 – 158). И вот критическое, но не подкрепленное ссылками замечание Пирсона (1930, с. 1):

*Кондорсе часто, а Лаплас иногда ошибались, потому что у них не было мыслей о корреляции.*

**15.** Лаплас (с. 342) вначале сослался на 7 лет наблюдений на Observatoire Royal, затем упомянул некоторые новые данные. Начиная по крайней мере с Ламарка, метеорологи узнали, что погода зависит от своего предыдущего состояния (Шейнин 1984b, § 5), и по этой причине возможно целесообразней было бы иметь дело не с разностями от одних суток к другим, а с внутрисуточными разностями.

Сохранение трех, а не двух значащих цифр (или, лучше, только одной) было крайне сомнительно, хотя и оправдано давней традицией.

**Основная литература**

Гнеденко и Шейнин (1978), Лаплас (1812), Пирсон (1978), Стиглер (1986), Шейнин (1973a; 1977), Эйзенхарт (1983)

## 4. Девятнадцатый век до 1809 г.

Мы описываем здесь работы предшественников Гаусса[1] и для полноты изложения рассматриваем задачи, которые ныне можно связать с МНКв. Сюда же мы включили работы Гаусса до 1809 г.

**4.1. Решение избыточных систем линейных уравнений.** Рассмотрим уравнения с двумя неизвестными

$$a_i x + b_i y + l_i = 0,\ i = 1, 2, \ldots, n. \qquad (1)$$

Самый ранний способ их решения состоял в разбивке системы на пары уравнений, решении каждой пары и осреднении полученных частных решений (§ 2.2 и начало § 2.5.1) Этот классический случай относился к определению (двух) параметров земного эллипсоида по градусным измерениям.

Пусть пара образована из уравнений $i$ и $j$. Ее решение, если оно существует и единственно, является точкой ($x_{ij}$; $y_{ij}$) пересечения двух соответствующих прямых:

$$x_{ij} = -A_{ij}/D_{ij},\ y_{ij} = -B_{ij}/D_{ij}, \qquad (2)$$

где в правых частях указаны соответствующие определители ($D_{ij}$ – определитель системы этой пары). Якоби (1841) и Бине независимо доказали, что решение системы (1) по МНКв является взвешенным средним из этих частных решений с весами $D_{ij}^2$:

$$x = \sum D_{ij}^2 x_{ij} / \sum D_{ij}^2,\ y = \sum D_{ij}^2 y_{ij} / \sum D_{ij}^2 .$$

Они даже установили, что это имеет место при любом числе неизвестных $k$, $2 \le k < n^2$. Впрочем, в XVIII в. частные решения (2) не взвешивались; в качестве окончательного решения принималось их обычное среднее арифметическое, т. е. центр тяжести точек ($x_{ij}$; $y_{ij}$). Заметим, что метод Бошковича решения систем (1), см. §2.5.1, также связан с частными решениями: в соответствии с (2) сумма остаточных свободных членов $v_i + v_j = 0$, что соответствует условию Бошковича (2.5).

Как было косвенно указано выше, случай $k > 2$ не представляет собой ничего нового, однако вычисления могут при этом оказаться слишком тягостными, ср. замечание Майера в § 2.2, и поэтому иногда решалась лишь небольшая часть всех возможных подсистем. Это заставляло выбирать лучшие сочетания уравнений, с чем Майер, видимо, неплохо справился. И снова следует упомянуть метод Бошковича, потому что он сводится к решению одной-единственной подсистемы (выбираемой по определенному правилу).

Пусть теперь заданы не уравнения, т. е. не прямые (1), а точки $(x_i;\ y_i)$ и требуется определить такую точку $(x;\ y)$, для которой

$$\sum\ [(x-x_i)^2+(y-y_i)^2\ ]=\min.$$

Решение очевидно: $x = \bar{x}$, $y = \bar{y}$. Эту задачу поставил Симпсон, который не указал ее происхождения и не опубликовал ее формулировку. Стиглер (1984, с. 619), обнаруживший ее среди рукописей Симпсона, полагает, что ее запись относится к 1760 г. Будь неизвестной не точка, а прямая, ныне можно было бы определить ее по МНКв.

**4.2. Старинная землеустроительная задача.** Направления *засечки* неизвестной точки $(x;\ y)$ с нескольких станций $(x_i;\ y_i)$, $i = 1, 2, \ldots, n$, нанесенные на планшет землеустроителя, не пересекаются в одной и той же точке, а образуют многоугольник ошибок. Требуется определить надежное положение точки. При аналитическом решении было бы естественно в наше время потребовать, чтобы координаты искомой точки удовлетворяли условию

$$\sum_{i=1}^{n} p_i\,(a_i x+b_i y+l_i)^2=\min,\ p_i=1/(a_i^2+b_i^2),$$

что сведется к решению избыточной линейной системы, т. е. к стандартной задаче МНКв. В те времена землеустроители, видимо, выбирали любую разумную точку внутри многоугольника ошибок.

**4.3. Хубер.** Он был швейцарским астрономом и математиком. К 1790 г. он опубликовал несколько астрономических статей и вскоре стал профессором математики Базельского университета. Мериан (1830, с. 148) указал, что Хубер обнаружил принцип наименьших квадратов еще до 1802 г.:

> *С ним случилось то, что происходит со многими учеными, живущими обособленно в небольших городах. Именно, он часто держал при себе многие полезные мысли ... Так, например, он уже раннее* [когда именно?] *самостоятельно открыл метод наименьших квадратов, который позднее стал известен по работам Гаусса и Лежандра*.

Комментаторы начиная с Вольфа (1858) соглашались с этим утверждением, однако Дутка (1990) обнаружил забытую статью (Шпис 1939), автор которой заключил противное. Исходя из неопубликованных вычислений Хубера и уравнивания “большой землемерной сети”, Шпис (с. 12) заметил, что Хубер был знаком с МНКв, но не изобрел его. Так, он (там же) процитировал Хубера:

*Вероятнейшие значения* $\Delta y$ *и* $\Delta x$ *могут быть определены по лежандровскому правилу наименьших квадратов.*

В 1805 г. Лежандр не упоминал вероятнейших значений.

**4.4. Лежандр.** Именно он (1805/2007, с. 74) определенно ввел принцип наименьших квадратов:

*Особенно важно следует поступать так, чтобы крайние ошибки без учета их знака были заключены в самые тесные как только возможно границы. Из всех принципов, которые могут быть предложены* [для решения избыточных систем линейных уравнений] *нет, как я полагаю, более точного или простого в применении, чем тот, который мы использовали в настоящей работе. Он состоит в том, чтобы привести к минимуму сумму квадратов ошибок* [точнее: остаточных свободных членов]. *Этот метод устанавливает своего рода равновесие между ошибками, которые, поскольку оно не позволяет преобладать крайним* [погрешностям]*, подходит для выявления состояния системы, наиболее приближающейся к истине.*

Рассмотрим теперь исходные уравнения (1) с $k$ неизвестными ($k < n$). Условие наименьших квадратов

$$W = [vv] = \min$$

приводит к

$$\partial W/2\partial x = \sum_{i=1}^{n} v_i \partial v_i/\partial x = 0,\ \partial W/2\partial y = \sum_{i=1}^{n} v_i \partial v_i/\partial y = 0, \ldots$$

или к

$$[av] = [aa]\hat{x} + [ab]\hat{y} + \ldots + [al] = 0,$$
$$[bv] = [ab]\hat{x} + [bb]\hat{y} + \ldots + [bl] = 0, \ldots \qquad (3)$$

Гаусс (1822, с. 141) мимоходом назвал эти уравнения нормальными и его термин укоренился. Новые неизвестные $\hat{x}$, $\hat{y}$, ... являются, разумеется, оценками исходных.

**4.5. Эдрейн.** Он (1808) обосновал нормальный закон и вывел принцип наименьших квадратов примерно в то же время, что и Гаусс и он также применил свои результаты к решению нескольких практических задач.

**4.5.1.** Нормальный закон (Дутка 1990). Вначале Эдрейн рассмотрел геодезическое измерение линий $a$ и $b$ с ошибками $x$ и $y$ соответственно и предположил, что

$$x/a = y/b,\ x + y = c. \qquad (4;\ 5)$$

Уравнение (4), однако, означало, что ошибки действуют систематически, а уравнение (5) никак не обосновано. Молчаливо предполагая, что $x$ и $y$ независимы (что заведомо неверно), Эдрейн поставил условие: при соблюдении уравнений (4) и (5) искомая плотность должна выполнять требование

$$f(x; a) \cdot f(y; b) = \max,$$

так что

$$f'(x; a)dx/f(x; a) + f'(y; b)dy/f(y; b) = 0,$$
$$f'(x; a)/f(x; a) = f'(y; b)/f(y; b).$$

В простейшем случае уравнению (5) соответствует

$f'(x; a)/f(x; a) = mx/a$ и т. д.

И здесь, и ниже Эдрейн дифференцирует по неустановленному аргументу. Рассматривая теперь линейное измерение одной линии, Эдрейн теперь полагает, что ошибки $x$ и $y$, направленные вдоль соответствующих осей координат, независимы (также неясно почему) и что

$$x^2 + y^2 = r^2. \quad (6)$$

Тогда, при соблюдении этого необоснованного условия, должно выполняться требование

$$f(x) \cdot f(y) = \max, \; W = f(x) \cdot f(y) - \lambda(x^2 + y^2) = \max,$$
$$f'(x) \cdot f(y) - 2\lambda x = 0, f(x) \cdot f'(y) - 2\lambda y = 0,$$
$f'(x)/xf(x) = f'(y)/yf(y) = c$ и т. д.

Доказательства Эдрейна, конечно же, никуда не годятся, но идея о том, что ошибки подчиняются некоторому закону распределения, который определяется в соответствии со свойствами ошибок (неверно, правда, установленными), совпадают с подходом Гаусса (§ 5.1.2). Более того. Второй вывод Эдрейна повторил Дж. Гершель и другие ученые вплоть до Максвелла. Кац (1939) и Линник (1952) вновь независимо рассмотрели его и ослабили предположения Максвелла (т. е. Эдрейна) о независимости.

Эдрейн также выписал совместное распределение двух независимых ошибок и указал, что соответствующие изолинии вероятностей являются эллипсами, – эллипсами ошибок, как они были позднее названы в теории ошибок[3].

**4.5.2.** Принцип среднего арифметического. Он немедленно следует из нормального закона: вероятность ошибок *a, b, c*, …, совершенных при измерении некоторой константы, минимальна, если

$$(x - a)^2 + (x - b)^2 + (x - c)^2 + \ldots = \min, \quad (7)$$
$$x = (a + b + c + \ldots)/n.$$

В общем случае, при уравнивании косвенных наблюдений, условие (7) становится принципом наименьших квадратов[4], ср.
§ 5.1.3. В отличие от Гаусса, Эдрейну не пришлось постулировать принцип среднего арифметического, но только ввиду неудовлетворительного вывода нормального закона.

**4.5.3.** Землеустроительная задача. Буссольный ход (отрезки ломаной) $A_1 - A_2 - \ldots - A_{n+1}$ с измеренными сторонами $a_1, a_2, \ldots, a_n$ (с погрешностями $\Delta a_1, \Delta a_2, \ldots, \Delta a_n$) и азимутами $\alpha_1, \alpha_2, \ldots, \alpha_n$ сторон (с погрешностями $\Delta\alpha_1, \Delta\alpha_2, \ldots, \Delta\alpha_n$) проложен между станциями с известными координатами $A_1$ $(x_1; y_1)$, $A_{n+1}$ $(x_{n+1}; y_{n+1})$[5]. Ошибки в положении вершин хода (концов отрезков ломаной) окажутся равными
**Рисунок**

$a_1\Delta\alpha_1, a_2\Delta\alpha_2, \ldots, a_{n-1}\Delta\alpha_{n-1}$.

Пусть $x_i - x_{i-1} = \Delta x_{i-1}$, $y_i - y_{i-1} = \Delta y_{i-1}$, $i = 2, 3, \ldots, n + 1$,

а погрешности этих приращений координат, $\xi_{i-1}$ и $\eta_{i-1}$, будут

$$\xi_{i-1} = \Delta\alpha_i \sin\alpha_i + a_i\Delta\alpha_i \cos\alpha_i,\ \eta_{i-1} = \Delta\alpha_i \cos\alpha_i - a_i\Delta\alpha_i \sin\alpha_i$$

при

$$\xi_1 + \xi_2 + \ldots + \xi_n = u,\ \eta_1 + \eta_2 + \ldots + \eta_n = v, \qquad (8)$$

где $u$ и $v$– соответствующие составляющие линейной невязки хода, $\sqrt{u^2 + v^2}$.

Именно эта задача была для Эдрейна исходной. Он уравнял ход, поставив условие: при выполнении уравнений (8) должно соблюдаться

$$(\Delta a_1)^2/a_1 + (\Delta a_2)^2/a_2 + \ldots + (\Delta a_n)^2/a_n +$$
$$[(a_1\Delta\alpha_1)^2/a_1 + (a_2\Delta\alpha_2)^2/a_2 + \ldots + (a_n\Delta\alpha_n)^2/a_n]/p^2 = \min \qquad (9)$$

и при $p = 1$ получил

$$\mu = u/(a_1 + a_1 + \ldots + a_n),\ \lambda = v/(a_1 + a_1 + \ldots + a_n),$$
$$\xi_i = \mu a_i,\ \eta_i = \lambda a_i. \qquad (10)$$

Формулы (10) применяются до сих пор (Чеботарев 1955, § 164), но условие (9) соответствуют исправлению хода за систематические, а не за случайные ошибки: Эдрейн принял веса сторон хода пропорциональными их длинам, а не корням квадратным из них, но заметим, что условие (9) сформулировано надлежащим образом, т. е. по отношению к непосредственно измеренным величинам, а не к $\xi_i$ и $\eta_i$, которые не независимы друг от друга.

**4.5.4.** Через десять лет[6] Эдрейн (1818a; 1818b) опубликовал две статьи, в одной из которых (1818a) применил принцип наименьших квадратов к изучению фигуры Земли, а точнее, – к уравниванию маятниковых наблюдений. Уже в 1808 г. он заявил, что исключил эту тему ввиду нехватки места, и за прошедшее после этого время подобное исследование, первое из вышедших в свет с применением принципа наименьших квадратов, опубликовал Био (1811, с. 167 – 169). Тем не менее работа Эдрейна интересна, поскольку он обнаружил две ошибки в уравнительных вычислениях Лапласа (*Неб. мех.*, т. 2, 1798, § 42 книги 3) по методу Бошковича.

Обозначим полуоси земного эллипсоида вращения через $a$ и b, $a > b$. Тогда его сжатие будет равно $\alpha = (a - b)/a$. Эдрейн, однако, принял сжатие в виде $\alpha_1 = (a - b)/b$ и получил 1/319, см. также ниже.

Во второй статье Эдрейн (1818b) вычислил радиус шарообразной Земли

$$r = (2a + b)/3 = 3959.36\text{миль}.$$

При $\alpha_1 = 1/319$ и 1*м* = 39.370113*дюйма* это равносильно $a$ = 6378.629*км*, в хорошем соответствии с эллипсоидом Красовского ($a$ = 6378.245*км*, вывод 1940 г.). Эдрейн, однако, не пояснил своего определения $r$, а при переходе к сжатию $\alpha = 1/298.3$ эллипсоида Красовского он получил бы $a$ = 6379.094*км*, – тоже не так плохо.

Первым европейским автором, который заметил статью Эдрейна 1808 г., был, видимо, К. Аббе (1871), а вот его последующие работы (в которых он, разумеется, сослался на первоначальную) стали более или менее известны, по крайней мере на какое-то время; много позже Штрассер (1957) не упомянул их.

Вот выдержка из письма Ольберса Гауссу 24.2.1819 (1900/1976, №1, с. 711):

> *Насколько я смог заключить из одной ... статьи, некий американец приписывает себе изобретение метода наименьших квадратов. Он ссылается на свою Алгебру* [?], *вышедшую еще в 1808 г.*

Гаусс на это ничего не ответил.

**4.6. Гаусс**

**4.6.1.** Принцип наименьших квадратов. Гаусс (1809a, с. 150; 1809b, § 186) указал, что применял его с 1794 или 1795 г. Во втором случае он упомянул "наш принцип", а позднее, в письме Лапласу 30.1.1812, *Werke*, том 10/1, с. 373 – 374) пояснил:

> *Я применял метод* [принцип] *наименьших квадратов с 1795 г. ... Но я начал часто применять этот метод лишь с 1802 г. и с тех пор применяю его, можно сказать, ежедневно в астрономических вычислениях* [орбит] *малых планет. ... Я не спешил публиковать изолированный отрывок, и Лежандр меня опередил. ... Я не думал. что г-н Лежандр может так высоко ценить идею столь простую, что следовало бы скорее удивляться, что ее не* [опубликовали] *сто лет назад ... Но я верю, что все, знающие меня, поверят мне на слово, так же, как*

*я поверил бы от всего сердца, скажи Лежандр, что он владел этим методом до 1795 г.*

В 1798 г. Гаусс (там же, с. 533) записал в своем *Дневнике* знаменитую фразу: “Защитил исчисление вероятностей от Лапласа” и прокомментировал эту запись в письмах 24.3.1807 и 24.1.1812 Ольберсу (1900/1976, №1, с. 329 и 493 – 494), см. § 4.6.2. В частности, раннее применение Гауссом МНКв косвенно подтвердил фон Цах (1813, с. 98 прим.):

*Прославленный д-р Гаусс владел этим методом с 1795 г. и с выгодой применил его при определении эллиптических орбит четырех новых* [малых] *планет, что усматривается из его замечательной работы* [*Теории движения*].

Оговоримся: из указанного сочинения этого всё-таки не следует. Но вот Жерарди (1977, с. 19, прим. 16), основываясь на архивных источниках, обнаружил, что Гаусс, который в 1802 – 1807 гг. участвовал в топографических работах (частично для своего собственного удовольствия), применил этот метод не позднее, чем в 1803 г. Имеется и много других случаев, в которых Гаусс вполне мог применять МНКв хотя бы для предварительных пробных вычислений или прикидок, притом что для него этот метод не был жесткой процедурой, см. § 2.2. Кроме того, возможные ошибки в исходных данных или неизвестный способ взвешивания наблюдений могли сделать подтверждение невозможным. Наконец, нельзя сбрасывать со счетов мнение современников Гаусса (например, того же фон Цаха), которые единодушно подтверждали его утверждение, сформулированное в письме Лапласу (см. выше).

Гаусс не только применял принцип наименьших квадратов, но и сообщил о своем открытии коллегам. Среди них мы обнаружили Бесселя (1832, с. 27) и Вольфганга Больяй, отца одного из открывателей неевклидовой геометрии, Яноша Больяй (письма Больяй от Сарториуса фон Вальтерсхаузена 12 и 28.8.1856, см. переписку Гаусса 1899/1987, с. 157 – 159), Ольберс же был в этом смысле хорошо известен и раньше.

Выражение Гаусса “наш принцип” (см. выше) возмутило Лежандра, который в письме Гауссу 31.5.1809 (*Werke*, том 10/1, с. 380) разумно заявил, что изобретение принадлежит тому, кто первым опубликовал его. Гаусс ничего не ответил, и Лежандр (1820, с. 79 – 80) обрушился на него с резкой критикой. Вот убедительное мнение Мея (1972, с. 309) об этом эпизоде:

*Гаусс очень заботился о своем приоритете ... Но для него это означало первым изобрести, а не опубликовать; и ему было достаточно устанавливать даты по личным записям, переписке, загадочным замечаниям в своих публикациях ... Намеренно или нет, он этим поведением сохранял преимущество тайны без потери приоритета в глазах последующих поколений.*

И, как бы продолжая мысль Мея, Бирман (1966, с. 18) по существу указал, что будущие поколения щадят только гауссов[7].

Мей (1972, с. 299) также заметил, что Гаусс “видимо” пришел к принципу наименьших квадратов при уравнивании приближенных вычислений квадратных корней и при “поиске закономерностей в распределении простых чисел”.

В связи с его предыдущим высказыванием мы сошлемся на старинную задачу Потенота, – на определение координат пункта (*D*) по измерению с него углов *ADB* и *BDC* между тремя данными станциями *A, B, C*. Вычисление оказывается невозможным, если все 4 точки располагаются на одной и той же окружности. Отвечая своему ученику, Герлингу, 24.10.1840, Гаусс ( 1927/1975, с. 615) разъяснил, как именно можно проще всего определить, что задача не имеет решения, но (с. 617) попросил своего корреспондента “пока” не разглашать его мысли, к которой он пришел примерно полвека назад (!), потому что он сам хотел бы это сделать и притом сказать [возможно] больше о применении комплексных величин. В противном же случае у него, Гаусса, отпадет желание возвращаться к этой задаче.

**4.6.2.** Нормальная плотность. Можно полагать, что Гаусс вывел ее в 1797 г. или несколько позже, быть может примерно так же, как он это описал в 1809 г. Он (1821, с. 143)

> *Впервые исследова*[л] *эту задачу в 1797 г. при помощи основных законов теории вероятностей* [и] *скоро убедился, что разыскание <u>вероятнейших</u> значений неизвестных величин было бы невозможно, не будь известна функция, которая представляет собой вероятность ошибок*.

Сам термин (нормальный закон), как заметил Краскл (1978), начал появляться с 1873 г. (Пирс), но окончательно его ввел Пирсон (1894).

**Примечания**

**1.** По поводу XVIII в. см. § 2.9.

**2.** Глейнсвик (1967) дополнительно указал, что взамен решения частных систем исходных уравнений можно решать соответствующие частные системы нормальных уравнений (и снова осреднять результаты). Более важно, что он также установил способ нахождения весов неизвестных.

**3.** Описывая историю теории корреляции, Карл Пирсон (1920) и Уокер (1928) рассмотрели результаты позднейших авторов (Плана, Браве), которые изучали погрешности положения точки на плоскости и в пространстве и заключили, что они не способствовали возникновению этой теории. И Пирсон (1920/1970, с. 185 – 187), и Уокер (с. 469) приписали аналогичные исследования Гауссу (1823b; 1828), который, однако, рассматривал нормальное распределение только в 1809 г. Более того, вопреки их утверждению, Гаусс не вводил членов, содержащих произведения переменных, а для сочинения 1828 г. они указали неверную дату.

**4.** Кулидж (1926) обнаружил, что в библиотеке Эдрейна был мемуар Лежандра (1805), – но когда он туда попал?

**5.** Эдрейн рассматривал замкнутый ход (замкнутый многоугольник), но различие между двумя вариантами не является существенным.

**6.** Хоган (1977) установил, что первый мемуар Эдрейна фактически появился в 1809 г.

7**.** Уместно привести высказывания Лапласа (1812, с. 353), а также самого Гаусса в письме Ольберсу 1824 г. (1900/1976, №1, с. 413), свидетельствующее о его отношении к Лежандру:

*Лежандр возымел простую идею рассматривать сумму квадратов ошибок наблюдений и приводить ее к минимуму, что непосредственно приводит к стольким же окончательным уравнениям, сколько элементов следует исправить. Этот ученый геометр первым опубликовал указанный метод, но надо отдать должное Гауссу, заметив, что за много лет до публикации Лежандра он постоянно пользовался той же идеей и сообщил о ней многим астрономам.*

*С негодованием и печалью я … прочел, что старика Лежандра, красу и гордость своей страны и своей эпохи, лишили пенсии.*

**Основная литература**

Гаусс (1809b), Дутка (1990), Лежандр (1805), Стиглер (1980; 1986), Шейнин (1993а)

## 5. Гаусс

О Гауссе до 1809 г. см. § 4.6.

**5.1.** ***Теория движения*** (1809). Книга появилась на латинском языке, потому что издатель не согласился на ее выход в свет на немецком. Вот что Ольберс указал в связи с этим в письме Гауссу 27.6.1809 (1900/1976, №1, с. 436):

*Вы были вполне правы, когда сказали мне, что при последовательном совершенствовании Вашего метода в его теперешнем виде он вряд ли подобен своему первоначальному состоянию. И переработка* [материала] *на латинский язык, насколько я могу припомнить по своему прежнему и лишь беглому просмотру немецкого текста, также намного усовершенствовала его.*

Немецкий текст не сохранился, а для нашей цели письмо Ольберса непонятно: имел ли он в виду также и МНКв или нет? Во всяком случае, ниже мы описываем содержание переводов с латинского текста.

**5.1.1.** Метод Бошковича (наш § 2.5.1). Пусть по этому методу уравниваются $n$ линейных уравнений с $m$ неизвестными. Тогда, как заметил Гаусс (§ 186), в точности $m$ остаточных свободных членов равны нулю[2]; в § 174 он посчитал это обстоятельство неблагоприятным, ср. его запись в своем *Дневнике* (наш § 4.6.1).

Утверждение Гаусса, которое легко доказать, означает, что ему была известна важная теорема линейного программирования. Далее. Поскольку из $n$ уравнений во внимание здесь принимается лишь $m$, метод Бошковича является “крайним” частным случаем метода сочетаний уравнений (§ 4.1), при котором вычислитель должен выделить эти $m$ уравнений в соответствии с определенными правилами (выделить должное базисное решение в задаче линейного программирования).

**5.1.2.** Нормальное распределение (§§ 175 – 177). Гаусс рассмотрел уравнения с $m$ неизвестными

$$M_i - V_i \equiv \varepsilon_i = -l_i + a_i x + b_i y + c_i z,\ i = 1, 2, \ldots, n. \qquad (1)$$

Так во всяком случае он указал в § 180, а в § 175 описал эти уравнения словесно[3], указав, что $V_i$ – функции неизвестных, а $M_i$ – результаты их “непосредственных наблюдений”. Это означает, что

$$V_i = -(a_i x + b_i y + c_i z),\ M_i = -l_i.$$

В § 177 Гаусс ввел ограничения:

$$V_1 = V_2 = \ldots = V_n = V = (1/n) \sum M_i.$$

Он не пояснил суть этого подхода, но, видимо, имел в виду случай одного неизвестного. Соответственно, он (§177) сформулировал принцип среднего арифметического[4]:

*Как аксиома должна быть принята гипотеза: если какая-нибудь величина будет определена из многих непосредственных наблюдений, произведенных при одинаковых обстоятельствах и с одинаковой тщательностью, то среднее арифметическое из всех наблюденных значений окажется наиболее вероятным значением, если не абсолютно точно, то по крайней мере очень близко к этому, так что всегда будет наиболее надежным придерживаться именно такого значения.*

Пусть теперь (§ 175) $\varphi(\varepsilon)$ – одновершинная и “в большинстве случаев” четная плотность ошибок наблюдения[5]. Тогда вероятность получения $n$ независимых наблюдений $l_i$ будет равна

$$\Omega = \varphi(\varepsilon_1)\, \varphi(\varepsilon_2) \ldots \varphi(\varepsilon_n),$$

где $\varepsilon_i$ – их ошибки, и, при равномерном априорном распределении каждого множества ($\varepsilon_1$; $\varepsilon_2$; … ; $\varepsilon_n$)[6], вероятнейшая система [оценок] неизвестных будет соответствовать наибольшему значению $\Omega$.

Сформулируем теперь задачу Гаусса для одного неизвестного. Даны наблюдения $x_1$, $x_2$, ..., $x_n$; принцип наибольшего правдоподобия приводит к

$$\varphi'(x_1 - \hat{x})/\varphi(x_1 - \hat{x}) + \varphi'(x_2 - \hat{x})/\varphi(x_2 - \hat{x}) + \ldots + \varphi'(x_n - \hat{x})/\varphi(x_n - \hat{x}) = 0, \qquad (2)$$

где $\hat{x}$ – оценка искомой константы. Требуется определить плотность $\varphi(x)$, при которой по принципу среднего арифметического $\hat{x} = \bar{x}$. Заметим, что Гаусс молчаливо предположил, что искомая функция дифференцируема.

Пусть

$$x_i = x_1 - nN,\ i = 1, 2, \ldots, n,$$

так что

$$x_1 + (x_2 + x_3 + \ldots + x_n) = x_1 + (n - 1)\, x_1 - n(n - 1)N,$$
$$N = (x_1 - \bar{x})/(n - 1),\ x_i - \bar{x} = -N,\ i = 2, 3, \ldots, n.$$

Уравнение (2) теперь запишется в виде

$$\varphi'(x_1 - \hat{x})/\varphi(x_1 - \hat{x}) = (1 - n)\varphi'(-N)/\varphi(-N),$$

$$\frac{\varphi'[N(n-1)]}{(1-n)\,\varphi[N(n-1)]} = \frac{\varphi'(-N)}{\varphi(-N)} = -\frac{\varphi'(N)}{\varphi(-N)}$$

(напомним: $\varphi(x)$ – четная функция),

$$\varphi'(x)/[x\varphi(x)] = \mathrm{Const},$$

$$\varphi(x) = (h/\sqrt{\pi})\exp(-h^2x^2),\ h > 0, \qquad (3)$$

где $h$ – мера точности наблюдений (gradus praecisionis, § 178)[7].

**5.1.3.** Принцип наименьших квадратов (§ 179) следует немедленно: вероятность получить $(\varepsilon_1; \varepsilon_2; \ldots; \varepsilon_n)$, см. уравнения (1), максимальна, если $[\varepsilon\varepsilon] = \min$. Этот принцип, как Гаусс заметил, “должен считаться за аксиому”.

Гельмерт (§ 7.6) и Мерримен (1877, с. 165) указали, по существу, что Гаусс после вывода плотности ошибок не стал различать их от остаточных свободных членов уравнений, а Чубер (§ 3.6.3) определил плотность этих последних при одном неизвестном в случае нормальных распределений $\varepsilon_i$.

**5.1.4.** Точность среднего арифметического. В § 173 Гаусс ввел формулу для вычисления точности среднего арифметического и обосновал ее в § 181[8]. Пусть (§ 181) неизвестная постоянная $x$ определяется из уравнений

$$a_ix = m_i + \varepsilon_i,\ i = 1, 2, \ldots, n.$$

Полагая, что ошибки наблюдений имеют плотность (3), он, видимо, в соответствии с § 180 установил, что

$$P(\varepsilon_1; \varepsilon_2; \ldots; \varepsilon_n) \sim \exp\{-h^2(x^2[aa] - 2x[am] + [mm])\}$$

и что поэтому оценка $\hat{x}$ искомого $x$ по МНКв равна

$$\hat{x} = [am]/[aa].$$

Гаусс также заметил, что

$$P \sim \exp \{-h^2([aa]\,(x^2 - 2x\hat{x} + [mm]/[aa])\},$$

так что мера точности оценки $\hat{x}$ оказывается в $\sqrt{[aa]}$ раз больше, чем для отдельного наблюдения $x = m$, при котором $a = 1$.

**5.1.5.** Точность случайной суммы. В рукописном примечании к §183 Гаусс заметил, что при $x = a + b + c + \ldots$

$$h_x = [(1/h_a^2) + (1/h_b^2) + (1/h_c^2) + \ldots]^{-1/2}.$$

Он, очевидно, имел в виду нормально распределенные величины *a, b, c*, … и, следовательно, знал, что сумма нескольких нормальных законов снова нормальна, ср. § 3.7.2 и 6.1.1d. Это примечание было найдено в принадлежавшем Гауссу экземпляре *Теории движения* и перепечатано в его *Werke* (также и в русском переводе соответствующего параграфа этой книги).

**5.1.6.** Точность [оценок] неизвестных (§ 182). Наше описание частично основано на позднейшей работе Гаусса (1811, § 13). Пусть начальные уравнения имеют вид

$$a_i x_1 + b_i x_2 + \ldots + s_i x_k + l_i = v_i,\ i = 1, 2, \ldots, n,$$

так что нормальные уравнения будут

$$\begin{aligned} \xi_1 &= [aa]\hat{x}_1 + [ab]\hat{x}_2 + \ldots + [as]\hat{x}_k + [al] = 0,\\ \xi_2 &= [ab]\hat{x}_1 + [bb]\hat{x}_2 + \ldots + [bs]\hat{x}_k + [bl] = 0, \ldots, \qquad (4)\\ \xi_k &= [as]\hat{x}_1 + [bs]\hat{x}_2 + \ldots + [ss]\hat{x}_k + [sl] = 0. \end{aligned}$$

При последовательном исключении неизвестных (метод Гаусса) окажется, что

$$\begin{aligned} \xi_1 &= [aa]\hat{x}_1 + [ab]\hat{x}_2 + \ldots + [as]\hat{x}_k + [al] = 0,\\ \xi_2 &= [bb1]\hat{x}_2 + [bc1]\hat{x}_3 + \ldots + [bs1]\hat{x}_k + [bl1] = 0, \ldots, \quad (4a)\\ \xi_k &= [ss\,(k-1)]\hat{x}_k + [sl\,(k-1)] = 0. \end{aligned}$$

Из этой редуцированной системы нормальных уравнений можно вычислить $\hat{x}_k$, $\hat{x}_{k-1}$, …, $\hat{x}_1$, равно как и $[vv]$:

$$[vv] = \xi_1^2/[aa] + \xi_2^2/[bb1] + \ldots + \xi_k^2/[ss\,(k-1)] + [llk],$$

причем, при решении по МНКв $[vv] = [llk]$.

Далее (Гаусс 1809b, § 182)

$$P(x_1; x_2; \ldots x_k) \sim \int_{-\infty}^{\infty} \exp(-h^2[vv])\, dx_1 =$$

$$\int_{-\infty}^{\infty} \exp\{-h^2([vv] - \xi_1{}^2/[aa] + \xi_1{}^2/[aa])\}\, dx_1$$

и, аналогично вычислениям в § 181 (см. наш § 5.1.4), Гаусс получил

$$P(x_1; x_2; \ldots x_k) \sim (\sqrt{\pi[aa]}/h) \exp\{-h^2([vv] - \xi_1{}^2/[aa])\}.$$

Интегрируя по $x_2$, $x_3$, …, $x_{k-1}$, он заметил, что точность $x_k$ (правильнее, $\hat{x}_k$) в $\sqrt{[ss(k-1)]}$ раз выше точности непосредственных наблюдений, у которых $h = 1$. По поводу остальных неизвестных см. его §§ 183 – 184 и, более подробно, в последующем его сочинении (§ 5.3.5)[9].

**5.2.** ***Определение точности наблюдений*** (1816). Гаусс определял меру точности $h$, – параметр нормального закона (3), см. §§ 5.2.1 и 5.2.2, исходя из квадратов и более высоких степеней ошибок.

**5.2.1.** Пусть ошибки $m$ [независимых] наблюдений обозначены через α, β, γ, … Гаусс указал, что вероятнейшее значение $\hat{h}$ меры $h$ определяется из условия

$$h^m \exp[-h^2(\alpha^2 + \beta^2 + \gamma^2 + \ldots)] = \max,$$

откуда

$$\hat{h} = \{m/[2(\alpha^2 + \beta^2 + \gamma^2 + \ldots)]\}^{1/2}.$$

Заметим, что $\hat{h} = 1/(\mu\sqrt{2})$, где μ – средняя квадратическая ошибка наблюдения. Гаусс также установил, что

$$P(\hat{h} - \lambda \le h \le \hat{h} + \lambda) = \theta(\lambda\sqrt{m}/\hat{h}),\ \theta(t) = (2/\sqrt{\pi})\int_0^t \exp(-z^2)\, dz,$$

так что, для $P = 1/2$, $\lambda = \rho\hat{h}/\sqrt{m}$, где $\rho \approx 0.477$ – корень уравнения $\theta(t) = 1/2$.

Наконец, для распределения (3) $P(|x| \le \rho/h) = 1/2$ и, следовательно,

$$r = \rho/h \qquad (5)$$

есть вероятная ошибка, которую формально ввел Бессель (1816, с. 141 – 142).

**5.2.2.** Пусть

$$S_n = |\alpha|^n + |\beta|^n + |\gamma|^n + \ldots,\ K_n = \int_{-\infty}^{\infty} x^n \varphi(x) dx.$$

Тогда для больших значений $m$

$$P(-\lambda \le S_n - mK_n \le \lambda) = \theta\,[\lambda/\sqrt{2m(K_{2n} - K_n^{\,2})}\,], \qquad (6)$$

где $mK_n$ – вероятнейшее значение $S_n$.

Эту формулу Гаусс[10] не обосновал, доказал ее Липшиц (1890), а Крамер (1946, § 28.2) указал, что предложение Гаусса является очевидным частным случаем ЦПТ.

Гаусс также вывел выражение для абсолютных моментов нормального закона

$$mK_n = \overline{S}_n = m\Pi(n-1)/2]/(h^n\sqrt{\pi}),\ \Pi(x) = \Gamma(x+1),$$

так что $h$, а потому и $r$, см. формулу (5), могли быть оценены по $S_n$, а кроме того могли быть вычислены вероятные интервалы для $r$. Сравнивая их для различных значений $n$, Гаусс заключил, что наилучшая оценка $r$ достигается при $n = 2$.

**5.2.3.** Асимптотическое распределение хи-квадрат. Пусть $1/(h\sqrt{2}) = \alpha$ и $n = 2$. Тогда

$$\sqrt{m(K_4 - K_2^{\,2})} = \alpha^2\sqrt{2m}$$

и, см. формулу (6),

$$P(-\lambda \le S_2 - mK_2 \le \lambda) = \mathrm{N}(0;\ \alpha^2\sqrt{2m}\,),$$

так что $S_2$ подчиняется нормальному закону $\mathrm{N}(m\alpha^2;\ \alpha^2\sqrt{2m}\,)$. Это и есть асимптотическое распределение хи-квадрат (Крамер 1946, § 20.2).

**5.3.** ***Теория комбинаций*** (1823b)

Гаусс снова устанавливает принцип наименьших квадратов, исходя теперь из принципа наибольшего веса (наименьшей дисперсии).

**5.3.1.** Случайные ошибки и нормальный закон. Гаусс (§ 1) выделил случайные (irregulares seu fortuiti) и систематические (constantes seu regulares) ошибки. Первые, не поддающиеся вычислению, вызваны либо несовершенством органов чувств или инструментов, либо внешними условиями (§§ 1 – 3). Понятие случайной величины еще не появилось и Гаусс, естественно, не упомянул его, см. § 2.6.

Гаусс (§ 4) предположил, что плотность ошибок наблюдений $\varphi(x)$ существует, одновершинна и, как и раньше (1809b, § 175), “в большинстве случаев” четна, так что (§ 5) их среднее значение равно нулю. Однако, очевидно понимая, что наблюдения не могут всегда обладать плотностями, которые отличаются друг от друга лишь параметрами (§ 5.1.2), он оставил свой прежний вывод универсальной плотности, т. е. оставил ограничение случайности нормальностью.

Вторая важная причина нового подхода упомянута в § 5.3.2. Наконец, мы полагаем, что Гаусс недолго был удовлетворен своим первым обоснованием МНКв. Действительно, его принцип среднего арифметического (§ 5.1.2) содержал оговорку, а выведенный принцип наименьших квадратов должен был “считаться за аксиому” (§ 5.1.3). Неудивительно, что Фрейденталь и Штейнер (1966, с. 177) заявили, что первое обоснование было “затейливым и малоубедительным”[11].

**5.3.2.** Мера точности. Соответственно, Гаусс (§ 6) ввел меру точности [дисперсию],

$$m^2 = \int_{-\infty}^{\infty} x^2 \varphi(x) dx, \qquad (7)$$

где $\varphi(x)$ была плотностью ошибок наблюдения. Выборочное значение дисперсии оказалось непараметрической оценкой. Он также указал, что разумно пользоваться интегральной мерой точности[12] и что среди подходящих функций $x$ простейшей является $x^2$, а в § 7 назвал *m средней ожидаемой ошибкой или просто средней ошибкой* (errorem medium metuendum, sive simpliciter errorem medium). Точность и вес (pondus) Гаусс там же определил как величины, обратно пропорциональные $m$ и $m^2$ соответственно.

**5.3.3.** Неравенство типа Бьенеме – Чебышева. Гаусс (§ 9) рассмотрел вероятность

$$\mu = P(|\xi| \le \lambda m) = \int_{-\lambda m}^{\lambda m} \varphi(x) dx,$$

где $\varphi(x)$ была [одновершинной] плотностью случайных ошибок $\xi$, для которых, естественно, $\text{E}\xi = 0$. В § 10 он доказал, что

$$\lambda \le \mu\sqrt{3} \text{ для } \mu \le 2/3; \ \lambda \le 2/[3\sqrt{1-\mu}\,] \text{ для } 2/3 \le \mu < 1.$$

Крамер (1946, § 15.7 и Упр. 4 к главам 15 – 20) привел современное доказательство этой “замечательной теоремы”.

**5.3.4.** Независимость. Гаусс (§ 18) был первым, кто определил это понятие. Ошибки $e_1$ и $e_2$ двух функций наблюдений, $V_1$ и $V_2$ (по контексту – линейных)[13], как он заявил, “не будут полностью независимы друг от друга”, если [хотя бы] одно наблюдение является их общим аргументом. Гаусс привел несколько примеров и в одном из них вычислил дисперсию линейной формы $W = [ce]$ независимых ошибок $e_i$:

$$m_W{}^2 = \sum c_i^2 m_i^2,$$

где $m_i^2$ – среднее значение квадрата $e_i$. Для двух функций таких ошибок, $V_1 = [\alpha e]$ и $V_2 = [\beta e]$,

$$\text{E}(V_1 V_2) = \sum \alpha_i \beta_i m_i^2 \ne 0.$$

**5.3.5.** Принцип наибольшего веса. Пусть исходные уравнения с $k$ неизвестными будут

$$a_i x + b_i y + c_i z = L_i = l_i + \varepsilon_i,\ i = 1, 2, \ldots, n,\ n > k.$$

Если их неравные веса, обозначенные через $p_i$, натуральные числа, то, для приведения уравнений к одному и тому же весу, каждое $i$-е из них следует выписать $p_i$ раз, или, при любых весах, умножить его на $\sqrt{p_i}$ . По условию наибольшего веса (см. ниже) оба случая приводят к обобщенному принципу наименьших квадратов $[pvv] = \min$.

Таким образом, вся теория может быть рассмотрена в предположении $p_i = \text{Const}$, что и будет сделано ниже.

Ошибки наблюдений предполагаются несмещенными[14], $\mathrm{E}\varepsilon_i = 0$, и требуется определить оценки неизвестных $\hat{x}$, $\hat{y}$, $\hat{z}$, ..., также несмещенные,

$$\mathrm{E}\,\hat{x} = x,\ \mathrm{E}\,\hat{y} = y,\ \mathrm{E}\,\hat{z} = z, \qquad (8a, b, c)$$

и обладающие наибольшими весами. Вывод Гаусса тяжеловесен и мы будем придерживаться изложения у Идельсона (1947, § 11). Рассмотрим первое неизвестное:

$$x = f\,(L_1;\ L_2;\ \ldots;\ L_n).$$

В соответствии с обычными ограничениями, принятыми в классической теории ошибок (несмещенность, линейность уравнений), неизвестные могут быть представлены линейной формой этих аргументов:

$$x = [\alpha l], \qquad (9a)$$

где $\alpha_i$ – некоторые множители. Затем, ввиду (8a)

$$\hat{x} = [\alpha l] \qquad (9b)$$

без свободного члена и, более того,

$$x = \sum \alpha_i\ (a_i x + b_i y + c_i z).$$

Это равенство должно быть тождественным, а потому

$$[a\alpha] = 1,\ [b\alpha] = [c\alpha] = 0. \qquad (10a, b, c)$$

Дисперсия $x$ очевидно равна

$$m_x^2 = m^2[\alpha\alpha] \qquad (11)$$

(см. § 5.3.4), где $m^2$ – среднее значение квадрата каждой ошибки наблюдения. Оценка $\hat{x}$, обладающая наибольшим весом, теперь может быть выведена по условию

$$W = [\alpha\alpha] - 2Q_{11}[a\alpha] - 2Q_{12}[b\alpha] - 2Q_{13}[c\alpha] = \max,$$

где $Q_{1j}$, $j = 1, 2, 3$ – множители Лагранжа. Поэтому

$$\alpha_i - a_iQ_{11} - b_iQ_{12} - c_iQ_{13} = 0. \qquad (12a)$$

Ввиду (10a)

$$[a\alpha] = [aa]Q_{11} + [ab]Q_{12} + [ac]Q_{13} = 1, \qquad (13a)$$
$$[b\alpha] = [ab]Q_{11} + [bb]Q_{12} + [bc]Q_{13} = 0, \qquad (13b)$$
$$[c\alpha] = [ac]Q_{11} + [bc]Q_{12} + [cc]Q_{13} = 0. \qquad (13c)$$

Уравнения (13a) – (13c)[15] определяют множители $Q_{1j}$, после чего может быть вычислены оптимальные значения $\alpha_i$, см. уравнение (12a). Далее, равенства (9a) и (12a) приводят к

$$\hat{x} = [al]Q_{11} + [bl]Q_{12} + [cl]Q_{13}, \qquad (14a)$$
$$\hat{y} = [al]Q_{12} + [bl]Q_{22} + [cl]Q_{23}, \qquad (14b)$$
$$\hat{z} = [aa]Q_{13} + [ab]Q_{23} + [ac]Q_{33}. \qquad (14c$$

Множители $Q_{12}$, $Q_{22}$ и $Q_{23}$ ($Q_{12} = Q_{21}$) получены из уравнений (13a) – (13c) со свободными членами (0; 1; 0) соответственно, а те же уравнения со свободными членами (0; 0; 1), учитывая, что
$Q_{23} = Q_{32}$, определят последние три множителя.

Осталось еще несколько шагов.

**5.3.5a.** Уравнения (14) приводят к

$$[aa]\hat{x} + [ab]\hat{y} + [ac]\hat{z} = [al],$$
$$[ab]\hat{x} + [bb]\hat{y} + [bc]\hat{z} = [bl], [ac]\hat{x} + [bc]\hat{y} + [cc]\hat{z} = [cl].$$

Таковы нормальные уравнения, соответствующие условию наименьших квадратов (§ 5.1.3), которое таким образом является следствием принципа наибольшего веса[16].

**5.3.5b.** При $m^2 = 1$ вес $x$ в силу равенства (11) оказывается равным $1/[\alpha\alpha]$. Его наибольшее значение будет, см. (12a),

$$1/([a\alpha]Q_{11} + [b\alpha]Q_{12} + [c\alpha]Q_{13}) = 1/Q_{11}. \qquad (15)$$

Таков, стало быть, вес $\hat{x}$.

**5.3.5c.** Аналогично соотношениям (9a), (12a), (10) и (11),

$$\hat{y} = [\beta l],\ \hat{z} = [\gamma l], \qquad (9c, d)$$
$$\beta_i = a_iQ_{21} + b_iQ_{22} + c_iQ_{23},\ \gamma_i = a_iQ_{31} + b_iQ_{32} + c_iQ_{33}, \quad (12b, c)$$

$[a\beta] = 0$, $[b\beta] = 1$, $[c\beta] = 0$, $[a\gamma] = 0$, $[b\gamma] = 0$, $[c\gamma] = 1$, (16a, b)

$$m_{\hat{y}}^{\ 2} = m^2\,[\beta\beta],\ m_{\hat{z}}^{\ 2} = m^2[\gamma\gamma],$$
$$p_{\hat{y}} = 1/[\beta\beta] = 1/Q_{22},\ p_{\hat{z}} = 1/[\gamma\gamma] = 1/Q_{33}.$$

**5.3.5d.** Значение $Q_{ij}$ для $i \neq j$. Из уравнений (12) ввиду (16) следует, что $Q_{12} = [\alpha\beta]$ и аналогично $Q_{13} = [\alpha\gamma]$, $Q_{23} = [\beta\gamma]$. В то же время соотношения (9) приводят к $E(\hat{x}\ \hat{y}) = [\alpha\beta]$ и т. д. Взаимосвязь между величинами $Q_{ij}$ и ковариациями Гауссу, естественно, не была известна.

**5.3.5e.** Дополнительные соотношения для последующего применения (см. § 5.3.8). Формулы (13) и отношения между $Q_{ii}$ и $Q_{ij}$ с одной стороны, и такими величинами, как $[\alpha\alpha]$, $[\beta\beta]$, $[\alpha\beta]$, … с другой, приводят к формулам

$$x = A + [\alpha\alpha]\xi + [\alpha\beta]\eta + [\alpha\gamma]\zeta,\ y = B + [\alpha\beta]\xi + [\beta\beta]\eta + [\beta\gamma]\zeta,$$
$$z = C + [\alpha\gamma]\xi + [\beta\gamma]\eta + [\gamma\gamma]\zeta, \qquad (17)$$

или, см. формулы (4), к

$$\hat{x} = A,\ \hat{y} = B,\ \hat{z} = C. \qquad (18)$$

В математической статистике несмещенные оценки с наименьшими дисперсиями, подчиняющиеся определенным аналитическим условиям, называются эффективными, так что оценки МНКв являются эффективными.

**5.3.6.** Функции оценок. Задана линейная функция

$$F = f\hat{x} + g\,\hat{y} + h\,\hat{z} \qquad (19)$$

оценок $\hat{x}, \hat{y}$, $\hat{z}$ неизвестных, см. систему (1), полученных по МНКв. Требуется определить ее вес. Аргументы этой функции не независимы (§ 5.3.4) и здесь требуется какой-то окольный путь. Разумеется, соотношения, подобные (9), приводят к

$$F = f[\alpha l] + g[\beta l] + h[\gamma l] = [\mu l],$$

так что $m_F^{\ 2} = m^2\,[\mu\mu]$, где $m^2$ – дисперсия каждого $l_i$. Гаусс (§ 29), однако, предпочел более практичный метод. Заметим, что эта задача естественно появляется в геодезии (см. ниже), чего он не указал.

Рассмотрим цепь треугольников, числом, скажем, 12. По одной из сторон каждого конечного треугольника, $AA_1$ и $BB_1$, базисы, чьи направления установлены астрономически; иными словами, они полностью закреплены. Углы треугольников измерены и должны быть уравнены так, чтобы невязки и треугольников, и те, которые выявляются при передаче по цепи направления и масштаба (азимутальное и линеаризированное базисное условия), исчезли[17].

Уравненные углы это оценки вида $\hat{x}$, $\hat{y}$, $\hat{z}$ в функции (19), а самой функцией может быть, например, длина какой-либо стороны слабейшего треугольника. Как и в § 5.3.5, мы следуем Идельсону (1947, § 13).

Пусть неизвестные множители $\mu_i$ удовлетворяют ограничениям

$$[\mu a] = f,\ [\mu b] = g,\ [\mu c] = h, \qquad (20)$$

где $a_i$, $b_i$, $c_i$ – коэффициенты исходных уравнений (1). Поскольку $\hat{x}$, $\hat{y}$, $\hat{z}$ обладают наибольшими весами, эти множители должны еще подчиняться условию

$$[\mu\mu] - 2Q_1[\mu a] - 2Q_2[\mu b] - 2Q_3[\mu c] = \max,$$

где $Q_1$, $Q_2$ и $Q_3$ – множители Лагранжа, так что

$$\mu_i = Q_1 a_i + Q_2 b_i + Q_3 c_i, \qquad (21)$$

$$[a\mu] = f = [aa]Q_1 + [ab]Q_2 + [ac]Q_3,$$
$$[b\mu] = g = [ab]Q_1 + [bb]Q_2 + [bc]Q_3,$$
$$[c\mu] = h = [ac]Q_1 + [bc]Q_2 + [cc]Q_3.$$

Приняв во внимание уравнения (13a) – (13c), нетрудно видеть, что

$$Q_1 = fQ_{11} + gQ_{12} + hQ_{13},\ Q_2 = fQ_{12} + gQ_{22} + hQ_{23},$$
$$Q_3 = fQ_{13} + gQ_{23} + hQ_{33}. \qquad (22)$$

Соотношение (21) теперь запишется в виде

$$\mu_i = f\,(Q_{11}\,a_i + Q_{12}\,b_i + Q_{13}\,c_i) + g\,(Q_{12}\,a_i + Q_{22}\,b_i + Q_{23}\,c_i) + h\,(Q_{13}\,a_i + Q_{23}\,b_i + Q_{33}\,c_i).$$

Ввиду уравнений (12)

$$\mu_i = f\alpha_i + g\beta_i + h\gamma_i$$

и поэтому, см. (20),

$$[\mu\mu] = Q_1[\mu a] + Q_2[\mu b] + Q_3[\mu c] = fQ_1 + gQ_2 + hQ_3,$$

или, в соответствии с уравнениями (22),

$$[\mu\mu] = f^2Q_{11} + g^2Q_{22} + h^2Q_{33} + 2fg\,Q_{12} + 2fh\,Q_{13} + 2gh\,Q_{23}.$$

**5.3.7.** Точность наблюдений (§§ 37 – 38).

**5.3.7a.** Лемма: $[vv] = [\varepsilon v]$. Доказательство: из уравнений

$$a_i x + b_i y + c_i z - l_i = v_i \qquad (23)$$

следует, что

$[av]\,\hat{x} + [bv]\,\hat{y} + [cv]\,\hat{z} - [lv] = -[lv] = [vv]$.

Но уравнения (1) приводят к $-[lv] = [\varepsilon v]$ и лемма доказана.

**5.3.7b.** Основная формула. Вычтем уравнение (23) из (1):

$$a_i(x - \hat{x}) + b_i(y - \hat{y}) + c_i(z - \hat{z}) = \varepsilon_i - v_i.$$

Имея в виду, что

$$x - \hat{x} = [\alpha L] - [\alpha l] = [\alpha\varepsilon],\ y - \hat{y} = [\beta\varepsilon],\ z - \hat{z} = [\gamma\varepsilon],$$

мы получим

$$[a\varepsilon]\,[\alpha\varepsilon] + [b\varepsilon]\,[\beta\varepsilon] + [c\varepsilon]\,[\gamma\varepsilon] = [\varepsilon\varepsilon] - [\varepsilon v] = [\varepsilon\varepsilon] - [vv].$$

Перейдем теперь к средним значениям обеих частей:

$$m^2\,([a\alpha] + [b\beta] + [c\gamma]) = nm^2 - \mathrm{E}[vv],\ m^2 = \mathrm{E}[vv]/(n - 3).$$

В общем случае $k$ неизвестных

$$m^2 = \mathrm{E}[vv]/(n - k), \qquad (24)$$

но поскольку невозможно установить $\mathrm{E}[vv]$, приходится вместо него подставлять само $[vv]$.

Цингер (1862, § 33) заметил, что в этой формуле уже *скрывается* МНКв.

Лаплас (1816) косвенно применил аналогичную формулу с $n$ вместо $(n - k)$ в знаменателе, хотя только для нормального распределения, Гаусс (1823a, с. 146) же, не называя никого, указал, что сотношение (24) следует предпочитать и по существу, и для поддержания “достоинства науки”.

**5.3.7c.** Границы. Не удовлетворившись этим результатом, Гаусс (§§ 39 – 40) определил границы для дисперсии $m^2$, т. е. для $\mathrm{D}m^2$. Его прямые вычисления несколько тягостны, но достаточно ясны и его окончательные границы были такими:

$$\frac{2(\nu_4 - 2m^4)}{(n-k)},\ \frac{\nu_4 - m^4}{n-k} + \frac{k\,(3m^4 - \nu_4)}{n}, \qquad (25)$$

где $\nu_4$ – четвертый момент ошибок. Словесно Гаусс добавил, что для нормального распределения ($\nu_4 = 3m^4$)

$$\mathrm{D}m^2 = 2m^4/(n - k). \qquad (26)$$

Более точно, выражения (25) и правая часть формулы (26) должны были включать не $m^2$, а неизвестную величину $\mathrm{E}\varepsilon_i^2$. Действительно, $\mathrm{E}m^2 = \mathrm{E}\varepsilon_i^2 \equiv s^2$, но $m^2 \neq \mathrm{E}\varepsilon_i^2$. Здесь $\varepsilon_i$ – истинная ошибка наблюдения $l_i$. Заметим также, что первое из двух выражений в (25) не всегда является нижней границей; их относительные расположения зависят от распределения ошибок. Первая граница ошибочна, см. § 7.7.

Колмогоров озаглавил один из разделов своей статьи (1946) *Догматическое изложение результатов Гаусса*. В нем он привел формулу (24), сразу же без символа ожидания, и формулу (26), – в его нумерации, формулы (XII) и (XIV), – но почему-то не упомянул, что последняя относилась только к нормальному распределению. Впоследствии он с соавторами (1947) исправил свою ошибку, см. § 7.7. Он указал еще, что формула (XII) является "просто определением", но, к сожалению, не развил этой мысли и никто из позднейших авторов не комментировал ее. Ср., однако, замечание Цингера в § 5.3.7b.

**5.3.8.** Включение нового наблюдения в проделанное уравнивание (§35). Исходные уравнения (1) приводят к отношениям (17). Пусть теперь появилось новое наблюдение

$$fx + gy + hz + k = v_{n+1} \qquad (27)$$

и требуется определить поправки к оценкам $\hat{x}$, $\hat{y}$, $\hat{z}$, а не уравнивать все $(n + 1)$ наблюдения заново.

Рассмотрим уравнение (27) как функцию $\hat{x}$, $\hat{y}$, $\hat{z}$. Обозначив $v_{n+1}$ просто через $v$, будем иметь

$$v = F\xi + G\eta + H\zeta + K,$$
$$F = f[\alpha\alpha] + g[\alpha\beta] + h[\alpha\gamma],\ G = f[\alpha\beta] + g[\beta\beta] + h[\beta\gamma],$$
$$H = f[\alpha\gamma] + g[\beta\gamma] + h[\gamma\gamma],\ K = k + Af + Bg + Ch.$$

Вес $v$ равен (см. § 5.3.6) $1/(Ff + Gg + Hh) \equiv 1/\omega$. Ввиду уравнений (4.3) и (4)

$$\xi = [av],\ \eta = [bv],\ \zeta = [cv],$$

где суммирование происходит от 1 до $n$. Новые значения этих величин оказываются равными

$$\xi^* = \xi + fv,\ \eta^* = \eta + gv,\ \zeta^* = \zeta + hv,$$
$$v = (F\xi^* + G\eta^* + H\zeta^* + K)/(1 + \omega).$$

Прежняя оценка $x$ теперь может быть записана в виде

$$\hat{x} = A + [\alpha\alpha](\xi^* - fv) + [\alpha\beta](\eta^* - gv) + [\alpha\gamma](\zeta^* - hv) =$$
$$A + [\alpha\alpha]\xi^* + [\alpha\beta]\eta^* + [\alpha\gamma]\zeta^* - F(F\xi^* + G\eta^* + H\zeta^*) + K/(1 + \omega).$$

Новое наблюдение может быть включено, если принять $\xi^* = \eta^* = \zeta^* = 0$ вместо $\xi = \eta = \zeta = 0$, ср. уравнения (4) и формулу (18), а прежняя оценка $\hat{x}$ теперь заменится ее новым значением

$$x^* = A - FK/(1 + \omega).$$

Ее вес, т. е. величина, обратная коэффициенту $\xi^*$, равен

$$1/[\alpha^*\alpha^*] = 1/\{[\alpha\alpha] - F^2(1 + \omega)\}$$

и аналогично

$$y^* = B - GK/(1 + \omega) \text{ с весом } 1/\{[\beta\beta] - G^2(1 + \omega)\} \text{ и т. д.}$$

Плакетт (1950) предложил современное доказательство формул Гаусса, относящихся к этой теме, или к *рекуррентному МНКв*, как он теперь называется (Спротт 1978, с. 185). Сославшись на другого автора, Спротт заметил, что этот метод стал исключительно важным при обработке данных.

**5.3.9.** Изменение веса наблюдения (§ 36). Пусть после уравнивания обнаружилось, что одно из них (т. е. одно из исходных уравнений) должно было иметь вес $p^*$, а не $p$. Требуется исправить оценки $\hat{x}$, $\hat{y}$, $\hat{z}$ не прибегая к новому уравниванию.

Гаусс заметил, что эта задача аналогична только что рассмотренной и привел необходимые формулы. Изменение веса равносильно добавлению нового наблюдения

$$fx + gy + hz + k = v_{n+1}, \text{ вес} = p^* - p,$$

или, иначе, к добавлению уравнения

$$(fx + gy + hz + k)\sqrt{p^* - p} = v_{n+1}\sqrt{p^* - p} = w_{n+1}, \text{ вес} = 1.$$

Задача действительна свелась к предыдущей.

**5.4. Практические соображения**

**5.4.1.** Число наблюдений. Сколько раз следует измерять углы треугольников при заданной степени точности результатов? Ни формула (24), ни теория ошибок в целом не учитывают наличия систематических ошибок и поэтому наблюдатель сможет установить достигнутую точность, да и то лишь частично, только после измерения всех трех углов каждого треугольника. К окончательной оценке приведет лишь измерение базисов и азимутов на обоих концах цепи и вычисление соответствующих невязок.

Тем не менее, если углы измерены при благоприятных условиях с соблюдением установленных правил для исключения систематических ошибок, надлежащим типом инструмента и определенным числом приемов, можно разумно надеяться, что заданная точность будет достигнута и что формула (24) это подтвердит.

Соответственно, к концу XIX в. или, возможно, несколько позднее по крайней мере в некоторых странах, например в Советском Союзе, см. также Бомфорд (1952/1971, с. 24), были введены жесткие правила для исполнения триангуляции высшего класса. Гауссу, однако, пришлось действовать иначе и он с этим успешно справился. Так (Шрейбер 1879, с. 141):

> *Из его* [Гаусса] *полевых журналов, которые находятся передо мной, скорее следует, что на каждой станции он наблюдал так долго, пока не убеждался, что каждый угол был измерен столько раз, сколько полагалось. И после этого он ... вводил*

*полученные значения направлений в уравнивание системы в качестве равноточных и независимых друг от друга*.

Добавим: вводил таким образом, несмотря на то, что углы измерялись резко отличными друг от друга количествами приемов (Гаусс, *Werke*, том 9, с. 278 – 281). Герлинг (1839, с. 166 – 167), бывший студент Гаусса, придерживался того же подхода и указал, что после некоторого числа приемов наблюдатель убеждается, что "всякое продолжение ... будет только напрасным ... И я именно так [т. е. соответственно] и поступал по примеру Гаусса".

И вот свидетельство Бесселя (1833, с. 464):

*Постоянные колебания в границах неизбежного несовершенства ... сообразно с самой сутью результатов, выводимых из наблюдений.*

Последующие авторы (Кларк 1880, с. 18 и 52; Дорси и Эйзенхарт 1969, с. 53) согласны в том, что число измерений не должно превышать определенной границы. Они, равно как и Курно (1843, §§ 130 и 138) и даже Бейес (Стиглер 1986, с. 94 – 95) обосновывали это утверждение наличием неизбежных [остаточных] систематических ошибок, и, добавим мы, некоторой зависимостью между отдельными наблюдениями.

**5.4.2.** Оценка действия случайных ошибок наблюдений. Наилучшей для этого является формула (24), дополненная верно установленными границами (§ 7.7). Она, однако, несколько ошибочна, поскольку [*vv*], будучи случайной величиной, не может всегда совпадать со своим математическим ожиданием E[*vv*], ср. §5.3.7b.

Видимо по этой причине Гаусс по крайней мере однажды, см. его письмо Герлингу 17.4.1844 (1927/1975, с. 687), вывел единое общее значение $m^2$ по результатам измерений на нескольких станциях, указав, что при небольшом числе наблюдений оценка их точности ненадежна. В других письмах 29.1.1847 Герлингу (там же, с. 744) и 19.4.1821 Бесселю (1880/1975, с. 382) он повторил свое указание.

Лаплас (§ 3.7.1) несомненно был того же мнения и современные авторы (Ку 1967, с. 309) согласны с этим.

**5.4.3.** Отбраковка уклоняющихся наблюдений. В каких случаях следут исключать наблюдение только потому, что оно уклоняется от других? Многие ученые, например Ламберт (§ 2.4.1) и Эйлер (§ 2.9), упоминали или даже пытались изучить целесообразность отбраковки. Видимо имея в виду оценку точности наблюдений, Лаплас (1818, с. 534) разумно заключил, что

*Чтобы успешно применять формулы вероятности к геодезическим наблюдениям, следует правдиво сообщать о всех тех обособленных результатах, которые были приняты, и не исключать никаких по той лишь причине, что они несколько* [!] *удалены от остальных*.

И вот Гаусс, в письме Ольберсу 3.5.1827 (*Werke*, том 8, с. 152 – 153), указал:

*Для успешного применения исчисления вероятностей к наблюдениям наивысшую важность всегда имеет обширное знание предмета. Если такого знания нет, то при не очень большом числе имеющихся наблюдений отбрасывание ввиду большого расхождения всегда сомнительно. Все отдельные составные части ошибок наблюдений, избежать которых вне нашей власти, имеют определенные границы, даже если мы и не в состоянии их точно указать. Существует очень много случаев, когда мы можем уверенно сказать, что происшедшая крупная ошибка лежит вне пределов возможности подобных ошибок и что по-видимому совершена чрезвычайная ошибка. Естественно, ее следует отбросить. Но поскольку можно представить, что эта ошибка возникла ввиду несчастливого стечения составных частей, ее не следует исключать. Иногда могут, конечно, иметься и такие случаи, когда сомнительно, следут ли причислять ошибки к первому или ко второму классу;* [тогда] *можно поступать как угодно, но принять себе за правило ничего не скрывать, чтобы другие могли по своему усмотрению считать также и по-другому. Числовые результаты будут, как ни считай, иметь равную пригодность, но, если слишком проворно отбрасывать наблюдения, возникнет опасность преувеличить их точность. Мне представляется, что это занятие более похоже на поступки в жизни, где редко или никогда не имеется математической строгости и где приходится поступать по наилучшему продуманному усмотрению*.

К отбраковке наблюдений мы вернемся в § 7.1, но добавить что-то к мнению Гаусса будет трудно.

**5.4.4.** Вычисления. При жизни Гаусса вычисления требовали больших усилий и решение обширных систем нормальных уравнений было особо тягостным. Гауссу удалось облегчить последнюю задачу введя метод последовательного исключения неизвестных, см. формулы (4а). Так, 14.5.1826 он сообщил Ольберсу (*Werke*, том 9, с. 320), что решил систему из 55 уравнений! Подробнее о его вычислениях см. Шейнин (1979, с. 53 – 54).

Вместе с тем, он по крайней мере однажды (письмо Герлингу 26.12.1823; там же, с. 278 – 281) решил систему нормальных уравнений итеративным методом, – тем его вариантом, который сейчас называется методом релаксации (Форсайт 1951; Шейнин 1963). Напомним также (§ 2.2), что иногда он применял приближенные методы и полагал, например (Гаусс 1809b, §185), что часто бывает достаточно вычислять коэффициенты нормальных уравнений приближенно.

**5.5. Обзор теории ошибок.** Этот обзор целесообразно поместить здесь, до описания работ Гельмерта (гл. 7-я).

**5.5.1.** Гаусс и Лаплас. В 1810 г. Лаплас (§ 3.3.2) рассмотрел уравнивание большого числа наблюдений, разделенных на группы, так что групповое среднее могло быть нормально распределено в соответствии с ЦПТ, и вывел принцип наименьших квадратов без всяких допущений о среднем арифметическом. Тем не менее, предположение большого числа наблюдений было ограничительным, а разделение на группы представляется искусственным, и во всяком случае Лаплас более не возвращался к этому подходу.

В 1811 г. Лаплас (§ 3.3.3) изучил уравнивание большого числа наблюдений с одним неизвестным ($x$). Он доказал, что оценка $\hat{x}$ этого неизвестного по МНКв обладает наименьшим абсолютным математическим ожиданием $\mathrm{E}|\hat{x}|$, а затем обобщил свое исследование на случай двух неизвестных. Не дисперсия, а погрешность типа $\mathrm{E}|\hat{x}|$ надолго оказалась его мерилом точности.

В 1812 г. Лаплас (§ 3.4.2) по сути повторил это исследование и кроме того (§ 3.4.3) вывел закон ошибок при помощи принципов наибольшего правдоподобия и среднего арифметического. Он, однако, предположил существование большого числа наблюдений и отыскивал плотность в виде $\exp[-\psi(x^2)]$.

Наконец, в 1818 г. Лаплас (§ 3.6.1), снова в соответствии с ЦПТ, признал нормальный закон как плотность распределения ошибок даже при небольшом числе наблюдений, поскольку посчитал (без достаточного обоснования), что ошибка наблюдения состоит из большого числа составляющих одного и того же порядка малости. Он тем самым выразил суть последующей теории элементарных ошибок[18] в соответствии с которой, в частности, принцип наименьших квадратов не зависел от принципа среднего арифметического.

Аналогично Гауссу (§ 5.3.2), Лаплас (1816, с. 499; 1818, с. 538; 1814/1999, с. 844, левый столбец) ввел вес наблюдения, равный, однако, параметру нормального закона

$$\varphi(x) = \sqrt{h/\pi}\ \exp(-hx^2).$$

Заметим, впрочем, что в последнем случае он добавил непонятное утверждение: "Вес среднего результата растет вместе с числом наблюдений, деленным [?] на число [неизвестных] элементов".

Лаплас (1816, с. 513) указал, что максимум веса соответствует наименьшей ошибке оценки неизвестного (наименьшему значению величин, подобных $\mathrm{E}|\hat{x}|$, см. выше) и таким образом сформулировал свой вариант принципа наибольшего веса, по существу связывая его с нормальным законом. В то же время, в 1818 г., он молчаливо признал дисперсию!

Гаусс неизменно изучал конечное число наблюдений. В 1809 г. он (§ 5.1.2) вывел нормальный закон, исходя из принципов среднего арифметического и наибольшего правдоподобия. В 1823 г. он (§ 5.3.2) выбрал дисперсию в качестве меры ошибок и доказал, что наименьшая дисперсия совместно с несмещенностью приводит к МНКв. Он кроме того оценил точность и наблюдений, и оценок неизвестных по МНКв, и линейных функций этих оценок, а также обратил особое внимание на методичность вычислений и ввел исключительно удобные обозначения. Именно он, а не Лаплас создал теорию ошибок как готовый рабочий инструмент, ср. мнение Субботина в § 0.4.

**5.5.2.** Восприятие теории ошибок. Сразу скажем, что результаты Гаусса оставались плохо известными. Хуже того: многие ученые, занимавшиеся обработкой наблюдений, были плохо знакомы с практически необходимыми формулами этого метода и вообще теории ошибок. Так, Айвори (1826 – 1830), которого Гаусс в письме Ольберсу 15.3.1827 (1909/1976, № 2, с. 475 – 476) назвал "проницательным математиком", опубликовал ряд статей об уравнивании маятниковых наблюдений, имея вначале лишь смутное представление о своей теме.

Еще менее известными были пояснения Гаусса, которыми он обосновывал свой подход и, в частности, его авторские сообщения (см. Библиографию) и появлялись совершенно ошибочные утверждения. Вот некоторые из них. Цингер (1862, с. 1): Лаплас будто бы предложил

*Строгое* [?] *и беспристрастное исследование. Из его анализа видно, что результаты способа наименьших квадратов получают более или менее значительную вероятность только при условии большого числа наблюдений, между тем как Гаусс старался на основании посторонних соображений придать этому способу безусловное значение* [ничего подобного]. *Если мы обратим внимание на то, что в законе больших чисел заключается вся сущность Теории случаев и что только при большом числе испытаний получают действительное фактическое значение все свойства случайных явлений, то нетрудно будет видеть справедливость лапласова вывода. При ограниченном же числе наблюдений мы вовсе не можем рассчитывать на взаимное уничтожение погрешностей и … всякое сочетание наблюдений может … повести столько же к увеличению погрешностей, сколько и к ослаблению их*.

На самом деле Гаусс (1809b, § 172; 1821, с. 142; 1823, § 6) указывал, что МНКв, хоть и целесообразен, но условен.

И вот просто бессмысленное утверждение (Некрасов, письмо Маркову 1913 г. (Шейнин 2005а, с. 256 – 257):

*Точки зрения Гаусса и Лапласа я различаю моментами относительно опыта. Первая точка зрения à posteriori, а вторая – à priori. Судить à posteriori удобнее, ибо данных больше, но эта точка зрения запаздывает, отстает, плетется за событием*.

Чебышев (1880/1936, с. 277 и след.) колебался между обоими обоснованиями, так и не сказав, что второе предпочтительнее. Его бывший студент Марков (1899), однако, решительно защитил принцип наименьшей дисперсии. Странным образом он (с. 246) тем не менее заявил, что МНКв удобен, но никакими другими достоинствами не обладает, – так для чего вообще надо было его обосновывать? В конце жизни Марков (1924, с. 323 прим.) указал, что остался при своем прежнем мнении о МНКв.

Что "Теория комбинаций" оставалась малоизвестной видно и по утверждению Фишера (1925/1990, с. 260):

*В тех случаях, где он подходящ, этот метод* [наименьших квадратов] *является специальным приложением метода наибольшего правдоподобия, из которого его можно вывести.*

Несколько десятилетий спустя Эйзенхарт (1964, с. 24) заметил, что существование второго обоснования МНКв

*Видимо по существу не известно почти никому из его американских пользователей за исключением студентов, изучающих повышенный курс математической статистики.*

О работе Маркова высказывались различные мнения. Во-первых, Нейман (1934, с. 595) ошибочно приписал ему второе гауссово обоснование МНКв, а Ф. Н. Дейвид и он (1938) усугубили эту ошибку, доказав *обобщенную теорему Маркова*, фактически установленную Гауссом. Неудивительно, что в 1950-е годы появилась на свет мистическая теорема Гаусса – Маркова, дожившая, как ни странно, до наших дней (Четтерджи 2003, с. 248 – 249). Ошибку Неймана заметили Плакетт (1949, с. 460) и Сил (1967/1970, с. 212).

Во-вторых, Линник и др. (1951, с. 637) заявили, что Марков по существу ввел понятия о несмещенных и эффективных оценках, но с таким же правом они могли бы здесь сослаться на Гаусса. Нейман (1934, с. 593), кстати, сделал верное заключение: Марков лишь четко сформулировал задачи о наилучших линейных оценках.

Добавим, что в 1910 г. сам Марков (Ондар 1977, с. 29) признал, что глава о МНКв в его руководстве, как ему *часто приходилось слышать*, изложена *недостаточно ясно*, Идельсон (1947, с. 101) же назвал ее *трудно написанной.*

Классическая формула (5.24) Гаусса для оценки точности наблюдений также описывалась неточно (Чебышев 1880/1936, с. 249 – 250) или даже вообще отрицалась (Бертран, § 7.10). Заметим еще, что Дедекинд (1860/1930, с. 99) отрицал эту формулу, потому что она приводила к неопределенности при числе наблюдений, равном числу неизвестных. Но так оно и должно было быть!

Обозначим наблюдения некоторой константы $a$ через $x_1, x_2, ..., x_n$ ($x_1 \le x_2 ... \le x_n$). Ученые древности измеряли надежность наблюдений их размахом ($x_n - x_1$), см. § 1.1, и эта практика сохранилась даже в XIX в. (Айвори 1830, с. 415). Точнее, кроме размаха естествоиспытатели и математики основывались либо на ($x_n - \bar{x}$) или ($\bar{x} - x_1$), либо на $(x_n - \bar{x})/\bar{x}$ или $(\bar{x} - x_1)/\bar{x}$, а в 1883 г. Рэйли (Мендоза 1991, с. 294) заявил, что “успех” наблюдений можно измерять “степенью соответствия чисел”.

Повторим (§ 1.1), что интервал $[x_n - x_1]$ вероятно возрастает с $n$, так что основываться на размахе сомнительно, а кроме того крайние наблюдения возможно искажены крупными ошибками; и, наконец, как оценивать косвенные наблюдения? Даже в 1955 г. Корнфельд, чью статью представил М. А. Леонтович, утверждал, что достоинство измерений достаточно измерять вероятностью

$$P(x_1 \le a \le x_n) = 1 - (1/2)^{n-1},$$

где *а* – искомая величина. Этот метод, если его можно так назвать, обоснован не более, чем применение размаха. Впервые его предложил Берви (1899), на которого Корнфельд не сослался.

**Примечания**

**1.** Вряд ли известно, что русский перевод *Теории движения* появился еще в 1861 г. Переводчиком был студент Московского университета И. М. Догель, но его фамилия не была указана на титульном листе. См. *Русская энц*., т. 5, с. 201; год издания неизвестен, но т. 1 вышел в 1911 г.

**2.** Он рассматривал только главное условие Бошковича. В §§ 188 – 189 Гаусс, как представляется, признал, что оно обеспечивает первое приближение.

**3.** Напомним (§0.1), что классическая теория ошибок имеет дело только с линейными уравнениями.

**4.** Позднее Гаусс (1845, с. 143) заметил, что для независимых наблюдений применение среднего арифметического является *в общем совершенно верным* и привело к *блестящим результатам* в естествознании. И вот мнение Гильберта (неопубликованная лекция 1905 г., см. Корри 1997, с. 161):

*Если для некоторой величины получены из наблюдений многие значения, то ее вероятнейшим значением является среднее арифметическое …*

**5.** Таково было гауссово косвенное определение случайной ошибки.

**6.** Принимать это в качестве независимого предположения не обязательно, “потому что его можно вывести из постулата среднего арифметического” (Уиттекер и Робинсон 1924, с. 219 прим.).

**7.** Определяя постоянную в формуле (3), Гаусс заметил, что интеграл от отрицательного квадрата показательной функции впервые вычислил Лаплас и повторил это в § 182. Лежандр, в письме Гауссу 1809 г. (§ 4.6), указал, что до Лапласа этот интеграл вычислил Эйлер.

**8.** Здесь, в *Теории движения*, Гаусс естественно ограничился случаем нормального распределения. Он обобщил свои формулы оценки точности в 1823 г. (§5.3.4).

**9.** Гаусс отметил возможность определения точности [оценок] всех неизвестных путем повторных решений нормальных уравнений. Позднее он (1823b, § 31) указал, что *некоторые вычислители* так и поступали, однако (1823a, с. 145) при большом числе неизвестных “таким искусственным способом ничего не удавалось выиграть”.

**10.** Фактически Гаусс применил абсолютные моменты и его выражение для $K_n$ формально неверно. Заметим еще, что $mK_n$ это среднее, а не вероятное значение $S_n$.

**11.** В 1831 г. Гаусс (*Werke*, том 8, с. 145 – 146) написал Энке, что *не без интереса* ознакомился с его попыткой обосновать принцип среднего арифметического детерминированными аксиомами. Многие авторы последовали за Энке, и Цох (1935) заключил, что, хотя успех и не был достигнут, этот принцип всё-таки может быть установлен без привлечения стохастических понятий. Содержательная сторона подобных исследований привела к появлению элементов теории инвариантных гипотез и оценок (Леман 1959/1979, гл. 6).

Чакрабарти (1989) попытался применить принцип среднего арифметического к термодинамике.

**12.** Он выразил то же мнение в некоторых своих письмах, особенно ясно в письме Бесселю в 1839 г. (*Werke*, том 8, с. 146 – 147):

> *То, что я впоследствии отказался от метафизики метода наименьших квадратов, приведенной в …* [в 1809 г.], *произошло главным образом по причине, о которой я сам публично не упоминал. Именно, я считаю во всех случаях менее важным отыскание такого значения неизвестной величины, вероятность которой максимальна, но всегда остается бесконечно малой, нежели того, с которым получаешь наименее невыгодную игру. Иными словами, если fa обозначает вероятность значения a для неизвестного x, то менее важно привести к максимуму fa, нежели к минимуму интеграл* $\int fx\, F(x-a)dx$, *распространенный на все возможные значения x, в котором за F берется функция всегда положительная и подходящим образом неизменно возрастающая при возрастании аргумента*.

Метафизику Гаусс возможно понимал как умозрительное начало, неполностью отражающую реальность (например, не имеющую места универсальность нормального распределения.

**13.** Без этого ограничения утверждение Гаусса было бы ошибочным. Действительно. В соответствии с теоремой Стьюдента – Фишера, как мы назовем ее, при нормальном распределении выборочные среднее и дисперсия независимы.

**14.** Несмещенность является одной из характеристик гауссовой теории ошибок, см. также ниже.

**15.** Заметим, что эти уравнения отличаются от нормальных только своими свободными членами.

**16.** Ярошенко (1893a; 1893b) попытался обосновать МНКв неравенством Бьенеме – Чебышева. Если задать некоторую вероятность $P$, то *теснейший* интервал $2\varepsilon$ для неравенства $P(|\xi - E\xi| < \varepsilon)$ имеет место при наименьшей дисперсии $D\xi$, и этот очевидный вывод позволяет выбрать оптимальные множители и *наилучшие* оценки (оценки МНКв) неизвестных, ср. мнение Лапласа в § 3.3.3. Можно сказать, что Ярошенко не сказал ничего нового, но по крайней мере включение неравенства Бьенеме – Чебышева явилось здесь интересным. Впрочем, Усов (1867) намного опередил Ярошенко.

**17.** В нашем примере 36 углов, 2 базиса и 2 азимута, т. е. всего 40 измерений, из которых необходимы только 24 + 1 + 1 = 26. Возникает, следовательно, 14 условий и неизвестными являются поправки углов. Базисное условие составляется по синусам некоторых измеренных углов, но, если поправки невелики, его можно линеаризировать.

Этот подход к уравниванию по методу *условных наблюдений*, описанный Гауссом (1828), практически очень важен, хотя никаких существенно новых идей в нем нет. Собственно уравнивание состоит в определении условного минимума суммы квадратов поправок с обычным применением множителей Лагранжа. Как ни странно, четкое пояснение сути указанного метода представил лишь Гельмерт (1872, с. 197).

В нашем примере сеть можно было бы уравнивать и по схеме *косвенных наблюдений* с 22 неизвестными координатами 11 промежуточных станций. Здесь, число избыточных измерений по-прежнему равно 36 – 22 = 14.

**18.** Первые соображения об элементарных ошибках обычно приписывают Хагену (1837), фактически же их применил уже Даниил Бернулли (§ 2.7.3). Никакой подобной теории (как ее называют) видимо не существует, но можно добавить, что при соответствующих условиях сочетание таких ошибок подчиняется нормальному распределению, как и заключили и Даниил Бернулли, и Хаген. При определении элементарных ошибок Хаген (1837/1867, с. 34) принял весьма стеснительные условия.

**Основная литература**

Гаусс (1957), Гнеденко и Шейнин (1978), Идельсон (1947), Мей (1972), Спротт (1978), Шейнин (1979; 1994b), Эйзенхарт (1964; 1978)

## 6. От Гаусса к Гельмерту и далее

Мы рассматриваем работы Гельмерта также в гл. 7-й.

В середине XIX в. несколько ученых внесли вклад в теорию ошибок (§ 6.1), которая также проникла в другие области естествознания (§ 6.2). В § 6.3 мы описываем идеи и результаты двух естествоиспытателей в метрологии и астрономии соответственно.- Нормальный закон снова и снова либо подтверждался, либо отвергался в качестве закона ошибок (§ 6.4), хотя ввиду определенных обстоятельств первое гауссово обоснование МНКв вовсе не отбрасывалось. Позднее, однако, начались попытки видоизменить нормальное распределение и тем самым установить новый универсальный закон ошибок (§ 6.5). Мы завершаем главу (§ 6.6), обсуждая взаимоотношения статистики и теории ошибок.

### 6.1. Некоторые новые работы

**6.1.1.** Бессель. Повторим (§§ 2.4 и 5.2.1), что он подхватил или заново после Ламберта ввел термин *теория ошибок* и формально определил вероятную ошибку. Иногда утверждается (но всякий раз без должного обоснования), что классическая формула Гаусса (5.24) для оценки точности наблюдений принадлежит Бесселю. Так, Фишер (1939, с. 3) заметил, что Гаусс [только] “доказал ее превосходство” (очевидно, относительно смещенной оценки Лапласа, см. § 5.3.7), тогда как Бессель впервые предложил ее для случая одного неизвестного. Позднее Фишер (1951, с. 39) повторил свое утверждение, притом вообще не упомянул Гаусса. Не претендуя на оригинальность, мы отрицаем авторство Бесселя.

Бессель, или во всяком случае его студент Розенбергер (1828), который сослался на него, описал уравнивание *косвенных наблюдений с условиями*, т. е. уравнений

$$a_i x + b_i y + c_i z + l_i = 0,\ i = 1, 2, \ldots, n, \qquad (1)$$

неизвестные в которых удовлетворяют условиям

$$\alpha_j x + \beta_j y + \gamma_j z + \delta_j = 0, j = 1, 2, \ldots, s, n - s > k.$$

Бессель (1838b/1961, с. 147) в свою очередь упомянул Розенбергера.

Уравнивание указанного вида может быть выполнено при помощи множителей Лагранжа и потому не представляет ничего существенно нового[1]; более того, его уже рассматривал Эдрейн (§ 4.5c). Зато Бессель оказался при этом первым, кто уравнял наблюдения, разбив их на группы.

Бессель (1838b, гл. 3) уравнял свою триангуляцию громоздким методом. Во-первых, он придал измеренным направлениям формально введенные веса, см. также его сложные вычиления (Бессель 1834). Во-вторых, он уравнял заодно станции и сеть в целом. Напомним, что Гаусс (§5.4.1) поступал совсем иначе и добавим, что метод Бесселя не прижился.

В своем основном сочинении по теории ошибок Бессель (1838a) попытался доказать ЦПТ, чтобы тем самым обосновать нормальное распределение ошибок наблюдений. Известно, что эта теорема поддалась только Чебышеву, да и то не вполне, но мемуар Бесселя содержал иные интересные идеи. В § 10 он перечислил 13 независимых источников ошибок, возникающих при измерении зенитных расстояний звезд. В § 2 он подметил существование составляющей ошибки измерения с антимодальной плотностью. Это еще не опровергало приложимость нормального распределения и вообще вряд ли многие читатели обратили внимание на подобный особый случай, поскольку Бессель посвятил свой мемуар обоснованию ЦПТ.

В §§ 1 – 2 Бессель привел два примера вычисления плотности распределения функции случайной величины, но допустил несколько ошибок (Куммель 1882, с. 177 – 180; Шейнин 2000b, с. 80 – 81).

Далее, Бессель (§ 7) доказал, что нормальный закон устойчив, т.е. что сумма двух (а потому и любого конечного числа) нормально распределенных случайных величин снова нормальна[2]. Это было известно Лапласу (§ 3.7.2) и Гауссу (§ 5.1.5), но они не выписали соответствующих формул, тем более не обосновали их.

В § 11 Бессель, исходя из работ Брадлея и своих собственных результатов, доказывал, что астрономические наблюдения действительно подчиняются нормальному закону. Тем не менее, и крупные, и мелкие погрешности произошли в приведенных им примерах более часто, а промежуточные – более редко, чем в соответствии с ним.

Бессель перепечатал данные Брадлея из своей прежней книги (1818), см. Шнейдер (1988, с. 279), в которой заявлял, что уклонения от нормальности были “слишком незначительны” и объяснил их недостаточным числом наблюдений. Никаких количественных правил он в то время еще не смог указать (их впервые предложил Пуассон лишь в 1837 г.), однако каждый из трех рядов Брадлея содержал не менее 300 наблюдений и вывод Бесселя никуда не годился. Комментаторы вплоть до Идельсона (1947, § 33.1) этого, однако, не заметили.

**6.1.2.** Коши (1853a; b) известен в теории ошибок в связи с *методом средних* решения систем уравнений вида (1). Изменим знаки (некоторых) уравнений так, чтобы все коэффициенты при $x$ оказались положительными, тогда $\sum a_i$ в суммарном уравнении

$$x\sum a_i + y\sum b_i + z\sum c_i + \sum l_i = 0 \qquad (2)$$

окажется сравнительно большой и $x$ можно будет определить из него, после чего исходная система запишется в виде

$$y\,[b_i - a_i\sum(b_i/a_i)] + z\,[c_i - a_i\sum(c_i/a_i)] + [l_i - a_i\sum(l_i/a_i)] = 0.$$

Такие же действия далее производятся по отношению к следующему неизвестному и т. д.

Фактически в правых частях уравнений (1) находятся ненулевые ошибки наблюдений $\varepsilon_i$ и уравнение (2) следовало бы подправить, записав его правую часть в виде $\sum\varepsilon_i$. После уравнивания эта сумма переходит в сумму остаточных свободных членов; при использовании МНКв она исчезает только если все коэффициенты хотя бы при одном неизвестном совпадают (если, например, $a_i$ = Const, см. § 3.5.2). Таким образом, метод средних не совпадает с МНКв.

Идельсон (1947, § 21) подробно описал метод средних и добавил. что он применялся в службах времени по крайней мере в двух странах. Линник (1958/1962, § 15.5) указал, что оценки, полученные по этому методу, являются несмещенными линейными формами свободных членов. Он также подсчитал их эффективность для случая одного и двух неизвестных[3].

**6.1.3.** Бьенеме (1852) указал, что оценки должны быть совместно эффективными. Теперь известно (Петров 1954, § 5), что таковы оценки МНКв, если и только если ошибки наблюдений взаимно независимы и нормально распределены и если притом выполнены некоторые общие аналитические условия[4]. Этот факт подчеркивает особую роль нормального распределения, но, видимо, редко упоминается в современной литературе.

Бьенеме (1853, с. 313) также заявил, что выбор средней квадратической ошибки в качестве меры точности вовсе не произволен, “как это полагал г-н Гаусс”; аддитивность, подобная

$$\mathrm{D}(\xi_1 + \xi_2 + \ldots + \xi_n) = \mathrm{D}\xi_1 + \mathrm{D}\xi_2 + \ldots + \mathrm{D}\xi_n, \qquad (3)$$

не имеет места для других разумных мер точности, например для четвертых моментов. Напомним, однако (§ 5.3.2), что Гаусс обосновывал среднюю квадратическую ошибку (и дисперсию) ее простотой и, что почти очевидно, Гаусс знал формулу (3), см. § 5.3.4.

Хейде и Сенета (1977) описали вклад Бьенеме в статистику вообще и подробно рассмотрели его мысли о совместной эффективности в своем § 4.3.

**6.2. Физика, химия, метеорология**

Как свидетельствует всё, сказанное до сих пор, теория ошибок возникла и развивалась в соответствии с требованиями астрономии и геодезии; особым исключением была *Фотометрия* Ламберта (§§ 2.4.1 – 2.4.3). Но уже Паукер (1819) посвятил элементарную книжечку приложению МНКв к физическим наблюдениям, а Штреккер (1846) применил его при определении атомных весов двух элементов[5]. Далее, Клаузиус описал МНКв в рукописи, см. Шнейдер (1975, с. 248, прим. 28), а Максвелл, в 1860 г., при изучении распределения скоростей молекул газа использовал соображения, взятые из теории ошибок (ср. § 6.4.1b) и, наконец, не позднее 1870 г. элементы теории ошибок появились в “практической физике” (Кольрауш). Мы теперь опишем общее положение в трех указанных областях естествознания.

**6.2.1.** Число наблюдений. Напомним (§ 2.6), что Бойль противопоставил качество и количество опытов, полагая, что второе не столь уж и важно[6]. И, несмотря на усилия Симпсона (там же), это мнение, или, возможно, традиция выбирать “лучшее” измерение сохранилось в естествознании. Джоуль (1849) определил 5 значений механического эквивалента теплоты из уравнений вида

$$a_i x = l_i,\ i = 1, 2, \ldots, 5,$$

но использовал лишь одно из них. Выбранное значение он установил по опыту, который состоял из наибольшего числа измерений, притом с наибольшими значениями $a_i$, так что погрешности измерения $l_i$ делились на наибольший делитель. Трудно отделаться от впечатления, что Джоуль заранее продумал всё это[7].

Мендоза (1991, с. 283) попытался выделить астрономию и геодезию, указав, что в этих отраслях естествознания наблюдатели производили

> *Большое число повторных измерений одного и того же простого количества, ... что не имело отношения к небольшому числу сложных измерений атомных весов или удельной теплоемкости.*

Можно предположить, что Мендоза в глаза не видел теодолита и уж, конечно, не присутствовал при измерении углов, см. по этому поводу Бессель (1838b, §15) и Бомфорд (1952/1958, § 1.2.1). Отличие между отраслями науки, если и было, то гораздо более тонкое. В астрономии и геодезии разрабатывались и совершенствовались методы наблюдения, а инструменты должны были юстироваться. В “простые количества” приходилось вводить несколько поправок, но остаточные систематические ошибки, например вызванные горизонтальной рефракцией, были не менее вредны, чем примеси в образцах в химии, притом же каждое “сложное измерение” в физике и химии видимо состояло из многих “простых” измерений (как у Джоуля).

Заметим, однако, что триангуляция в данном районе прокладывается только один раз, тогда как константы в физике, химии (и астрономии!) определяют в нескольких местах, возможно на протяжении многих лет.

**6.2.2.** Взвешивание наблюдений. Несколько наблюдений (например, *s*), проведенных в одних и тех же условиях, искажены систематическими ошибками быть может примерно в одинаковой степени, так что приписывать вес *s* их среднему арифметическому неразумно. Гораздо целесобразнее наблюдать при изменяющихся (но всё же благоприятных) условиях. Герлинг (1839, с. 14 – 15) пояснил, что [по крайней мере]

> *Те станции, визирование на которые в различное время дня, либо при жаре и покрытом небе, как мы опасались, происходило с отклонениями луча* [с фазами]*, наблюдались как можно чаще, также и при этих различных условиях ...*

Вряд ли Гаусс поступал иначе. Обратим теперь внимание на маятниковые наблюдения, по которым определялось сжатие земного эллипсоида вращения. Био (1829, с. 16 – 17) указал, что пункты наблюдения, расположенные примерно на одной и той же широте, можно легко заменить одной фиктивной станцией.

Действительно, период колебаний маятника заданной длины, а также и его длина при заданном периоде колебаний, зависят только от широты места. Но если пункты на самом деле близки друг к другу (т. е. если их долготы также примерно одни и те же), то все наблюдения возможно искажаются примерно одной и той же местной гравиметрической аномалией и вес осредненной станции в таких случаях вряд ли следует увеличивать по сравнению с весом отдельного пункта.

**6.2.3.** Зависимость между наблюдениями. Особые трудности в связи с зависимостью возникали в метеорологии[8], поскольку смежные значения данного метеорологического элемента явно зависимы. Вот интересная попытка обойти эту трудность (Ламон 1867, с. 245):

> *Я уже 30 лет назад начал* [при обработке данных] *заменять обычные метеорологические наблюдения разностями одновременных наблюдений* [произведенных в разных местах] *и теперь я убедился, что этот подход самый подходящий для построения метеорологии как математической дисциплины*.

Ламон (прим. 1839, с. 263) раннее даже утверждал. что одинединственный год изучения подобных разностей столь же надежен, как результаты *обычных* 30 лет, а в 1849 г. Кетле (Шейнин 1984b, с. 72, прим. 41), который весьма успешно занимался метеорологией, заметил, не ссылаясь ни на кого, что разности, подобные упомянутым Ламоном, имеют характер случайных ошибок.

**6.3. Метрология и астрономия.** Эти науки всё-таки в большей степени используют теорию ошибок, и здесь мы опишем работы Менделеева и Ньюкома соответственно. Первого мы уже упоминали в прим. 11 к гл. 3-й и в § 6.2.1, прим. 7, а о втором будем еще говорить в §§ 6.4.2 и 6.5.3.

**6.3.1.** Менделеев. С 1893 по 1907 гг. Менделеев был директором Главной палаты мер и весов, и обработкой наблюдений он занимался и как химик, и как метролог. Он (1887, с. 82) обращал особое внимание на доброкачественность измерений и возражал против объединения наблюдений, проведенных различными методами и в различных условиях, равно как и против их чрезмерного накопления.

Вот утверждение Менделеева (1872, с. 144) об определении соотношения между плотностью и давлением газа при уточнении закона Бойля – Мариотта:

> *Я предпочитаю сделать немногие, но точные и повторенные наблюдения при нескольких значительно разнящихся давлениях, чтобы по возможности не прибегать к способу наименьших квадратов. ... Умножение числа наблюдений при разнообразных давлениях, близких друг к другу, представляет не только многие затруднения для исследования, но и увеличивает погрешности вывода.*

Менделеев видимо имел в виду влияние систематических ошибок, которые могли зависеть от давления; заметим, что он здесь же рекомендовал получать три – четыре наблюдения при каждом изучаемом давлении и что нежелание обращаться к МНКв было безусловно связано с трудностями вычисления.

Вот утверждение Менделеева (1872, с. 144) об определении соотношения между плотностью и давлением газа при уточнении закона Бойля – Мариотта:

> *Я предпочитаю сделать немногие, но точные и повторенные наблюдения при нескольких значительно разнящихся давлениях, чтобы по возможности не прибегать к способу наименьших квадратов. ... Умножение числа наблюдений при разнообразных давлениях, близких друг к другу, представляет не только многие затруднения для исследования, но и увеличивает погрешности вывода.*

Менделеев видимо имел в виду влияние систематических ошибок, которые могли зависеть от давления; заметим, что он здесь же рекомендовал получать три – четыре наблюдения при каждом изучаемом давлении и что нежелание обращаться к МНКв было безусловно связано с трудностями вычисления.

В прим. 11 к гл. 3-й мы упомянули, что Менделеев недооценил медиану. Он (1875, с. 209 прим.) оставил еще одно неверное утверждение об этой оценке: если “вероятный вывод” (т. е. медиана) “совершенно согласен” со средним арифметическим, то

> *Погрешности следуют определенному закону, принятому гауссовской теорией вероятностей, т. е. наблюдения не содержат крупных случайных уклонений, а определяются неизбежными погрешностями наблюдений.*

Крупными уклонениями Менделеев, видимо, называл промахи, но вот гауссовской теории вероятностей просто не существует. Он продолжал:

*Я … пользуюсь иным приемом для оценки стройности ряда наблюдений, долженствующего давать тождественные числа, а именно распределяю все числа на 3, по возможности равные, разряда … среднее из среднего разряда считается вероятнейшим …, а если к нему близко среднее из остальных разрядов …, то наблюдения считаются стройными. Но в расчетах Главной палаты мы держимся обычных выводов Гаусса.*

Если наблюдения должны быть тождественны, то это, видимо, означает, что измеряется истинное значение чего-то (§ 6.5.1). Заметим, что определенным образом нормированное уклонение среднего арифметического от медианы признается в качестве меры асим-метрии распределения (Юл и Кендалл 1937/1958, с. 161).

В качестве основной меры точности Менделеев принимал вероятную ошибку, а допустимым расхождением между двумя средними он (1860, с. 46) полагал сумму их вероятных ошибок. Пусть указанные ошибки равны друг другу, тогда, для случая нормального распределения, их сумма равна $1.35m$ ($m$ – средняя квадратическая ошибка каждого среднего). С другой стороны, та же ошибка разности средних равна $m\sqrt{2}$. Таким образом оказывается, что исследуемая разность существенна по Менделееву, если она равна своей средней квадратической ошибке. Это требование, пожалуй, слишком жестко; сравнительно недавно рекомендовалось иное правило (Дорси и Эйзенхарт 1969, с. 50), в котором вероятные ошибки относились к единичным измерениям.

И вот соответствующее высказывание Маркова из малоизвестного источника (*Прения* 1903; Шейнин 1990b, с. 453 – 454), которое свидетельствует, что указанное выше правило было общепринято в естествознании, см. также § 6.3.2:

*Правило Ф. А. Бредихина, что для признания реальности величины вычисленной требуется, чтобы она по крайней мере в два раза превосходила свою вероятную погрешность, мне очень нравится. Я не знаю только, кто установил такое правило и признают ли его все опытные вычислители.*

Иными словами, разность между нулем и реальной ненулевой величиной должна превышать ее удвоенную вероятную ошибку.

**6.3.2.** Ньюком. Он обработал более 62 тысяч наблюдений Солнца и планет (Бенджамин 1910) и существенно исправил принятые до него значения астрономических констант. Для этого ему пришлось исследовать и сравнивать наблюдения, выполненные на главных обсерваториях мира (к чему, кстати, Менделеев, см. § 6.3.1, отнесся бы с сомнением), притом, пожалуй, без всяких вычислительных средств кроме таблиц логарифмов.

По крайней мере в одном случае Ньюком (1872) назначил веса отдельным астрономическим каталогам в зависимости от их систематических погрешностей (вряд ли обходясь без субъективных представлений), а на заключительной стадии уравнивания повторил объединение каталогов с весами, соответствующими теперь уже случайным ошибкам. Он неоднократно исследовал различные систематические ошибки и мы упомянем изучение изменений личного уравнения с величиной звезд (1896a; 1897c, с. 187).

Вечная проблема уклоняющихся наблюдений, разумеется, заботила его. Вначале он относился к ним с большим сомнением, затем, однако (1895, с. 186), стал в этом отношении терпимее. Если ряд наблюдений нельзя было характеризовать нормальным распределением, Ньюком предпочитал назначать меньший вес *дальним* результатам (1896b, с. 43), – к чему Менделеев (§ 6.3.1) снова отнесся бы с сомнением, – либо, при асимметричных рядах, выбирать медиану вместо среднего арифметического. На Курно (§ 3.2.1b, прим. 4) он, впрочем, не сослался, и в двух одновременно вышедших мемуарах (1897a; 1897b) назвал медиану двумя (!) другими, ныне забытыми терминами.

Как и Менделеев (§ 6.3.1), Ньюком полагал, что расхождение между двумя эмпирическими величинами существенно, если оно превышает сумму соответствующих вероятных ошибок. И всё-таки Ньюком несколько раз указывал, что выведенная им величина $a$ имела среднюю квадратическую ошибку $b$ даже при значительном превышении последней над первой, вплоть до случая $a = 0.05$ и $b = 0.92$ (1901, с. 9)!

Неоднократно применяя МНКв, Ньюком иногда отступал от правил. Так, следуя указанию Гаусса (§ 5.4.4) и примеру предшественника, он (1897a, с. 31) подчас вычислял коэффициенты нормальных уравнений приближенно; в другом случае он (1895, с. 52) посчитал, что малыми коэффициентами в системе нормальных уравнений можно пренебрегать, хотя и не указал никакого количественного правила.

Далее, Ньюком представлял себе, что зависимость между нормальными уравнениями может происходить ввиду накопления погрешностей при переходе к ним от исходных уравнений, и разумно решил, что в таких случаях следует вычислять с вдвое бóльшим числом значащих цифр. Именно так он (1867) поступил при исследовании вычислений казанского астронома Ковальского, который заметил, что из выведенных им четырех нормальных уравнений лишь два независимы. Сейчас известно, что плохо обусловленные исходные уравнения лучше решать без перехода к нормальным, например, методом последовательных приближений.

Отметим еще особое вычисление. Имея 89 исходных уравнений с пятью неизвестными, Ньюком (1874, с. 167) составил и решил нормальные уравнения, Но затем, вычислив остаточные свободные члены первых, он всё-таки решил их каким-то иным способом (представив лишь результаты обоих решений). Можно полагать, что он имел в виду по возможности исключить систематические погрешности, но как?

В теоретическом плане Ньюком (1895, с. 82; 1897b, с. 161) ошибался, хотя и ссылался раннее на второе гауссово обоснование МНКв, полагая, что этот метод неотделим от нормального распределения. Укажем еще на его неудачное рассуждение (Ньюком и Холден 1874, с. 270 – 271): для систематической ошибки $s$ и случайных ошибок $r_1$ и $r_2$ он специально доказывал, и притом лишь для нормального распределения, рассматривая соответствующий двумерный интеграл, что

$$E[(s + r_1)(s + r_2)] = s^2.$$

Остановимся, наконец, на исследовании колебания широт (Ньюком 1892). Было известно, что они вызываются движением полюса около некоторой точки, в основном по кривой, близкой к окружности, с периодом 1.2 года. Ньюком проверял выдвинутую в то время гипотезу о том, что колебания периодичны с периодом 1.17 года и предположил, что полюс движется равномерно по окружности. Некоторые его вычисления сомнительны (и недостаточно подробны, что характерно и для многих иных его работ), однако его вывод (гипотеза не опровергалась) оказался верным.

### 6.4. Нормальный закон

**6.4.1.** Успех. По нескольким причинам нормальный закон в течение многих десятилетий считался законом ошибок. Во-первых, его математическое обоснование в *Теории движения* (§ 5.1.2) представлялось изящным, так что авторы учебников обычно упускали гораздо более сложное обоснование МНКв в “Теории комбинаций” (§ 5.3). Во-вторых, нормальный закон укоренился в естествознании; в третьих, показательная функция отрицательного квадрата была удобна в обращении; в четвертых, нормальное распределение более или менее хорошо описывало разброс ошибок наблюдения (несомненно в силу ЦПТ, которая к тому же придавала ему почтенность). Наконец, в пятых, нормальный закон был устойчив (§ 6.1.1d). Остановимся на этих соображениях несколько подробнее.

**6.4.1a.** Вряд ли кто-либо последовал за Гельмертом (Гаусс 1887, Предисловие, с. IV – V), который рекомендовал описывать оба обоснования МНКв. Отметив, что второе из них было предпочтительнее, он добавил:

> *Однако, никто из тех, кто имеет такую возможность, не должен отказываться от дополнительных указаний и именно изучить также и первое гауссово представление, тем более, что сам Гаусс ... в своих лекциях ... всегда начинал с прежней разработки и лишь затем позволял себе обращаться к “Теории комбинаций” ...*

В своем собственном сочинении Гельмерт (1872/1907, Предисловие) заявил, что

> *При выводе основных формул* [он] *еще отчетливей* [чем в первом издании] *поставил на первое место простой принцип наименьших квадратов. Лишь после этого* [он] *показал, при каких условиях решенные таким образом проблемы (Aufgaben) обладают и другими свойствами, важнейшее из которых то,*

*что при определенных допущениях веса* [оценок] *неизвестных ... окажутся наибольшими. Таков простейший путь*.

**6.4.1b.** В 1860 г. Максвелл заявил, что в состоянии равновесия распределение скоростей газовых молекул нормально; в астрономии, с середины XIX в. до 1896 г. направления собственных движений звезд считались нормально распределенными (Шейнин 1984a, § 8.4) и сами астрономические наблюдения по свидетельству Бесселя (§ 6.1.1e), вовсе не убедительному, полагались таковыми же.

**6.4.1c.** Более того. Нормальная плотность в естествознании оказалась к тому же законом, управляющим погрещностями, “допущенными природой ” (Кетле 1853, с. 64 – 65). Так, после изучения некоторого антропометрического распределения для нескольких тысяч призывников Кетле (Стиглер 1986, с. 260 и след.) заявил, что уклонения от среднего подчинялись нормальной плотности. И уже в конце жизни, в 1873 г., он (Шейнин 1986, с. 313) утверждал, что нормальный закон “является одним из самых общих в живой природе”.

**6.4.1d.** Хоть нормальная кривая не может описывать некоторые наблюдения в естествознании, они и не считались случайными в “обычном” смысле. Вот, к примеру, замечание Кетле (1846, с. 168) об атмосферном давлении:

*Понижение уровня ртути относительно среднего в общем более значительно, чем его повышение. Случаи, при которых среднее не находится на равном расстоянии от крайних значений и кривая возможностей лишается своей симметрии, весьма часты; они заслуживают быть изученными, особенно потому, что отсутствие симметрии всегда вызывается более или менее странными причинами, влияние которых можно оценить*.

Другой автор (Мейер 1891, с. 32) заявил, что, поскольку соответствующие кривые плотности асимметричны, “Исчисление ошибок в принципе не применимо в метеорологии”. Он мог бы добавить, что, к примеру, разброс местной температуры воздуха в течение суток не является случайным. Но интересно, что Пирсон (1898) использовал данные Мейера для проверки применимости своей теории асимметричных кривых.

**6.4.1e.** Неудивительно, что многие ученые обратили внимание на всеобщую веру в нормальный закон. По этому поводу Пуанкаре (1896/1912, с. 171) повторил устное замечание Г. Липпмана: “Экспериментаторы воображают, что это – математическая теорема, а математики полагают, что это – экспериментальный факт”[9].

**6.4.2.** Противники нормального закона. Строго говоря, Бессель (§6.1.1e) был первым, кто подметил отклонения от нормального закона, но его слабое утверждение было забыто. Бьенеме (1853, с. 313) заявил, что показательная функция является “лишь весьма удобным средством приближения, которую можно заменить другими формулами”.

Первым действительным противником нормального закона оказался Ньюком (1886, с. 343). Он полагал “весьма особыми” те случаи, при которых ошибки подчиняются ему, а на с. 345 добавил, что в некоторых “классах важных наблюдений” доля крупных ошибок столь значительна, что “не было никакой возможности отделить наблюдения на нормальные и не нормальные” (обычные и необычные?). Ньюком также упомянул свое прежнее утверждение (1882, с. 382):

*Всякое собрание наблюдений прохождений Меркурия* [по диску Солнца] *обязательно является смесью наблюдений с различными вероятными ошибками* [ср. § 6.5.3] *и это стало ясно автору из его наблюдений ... 1878 г. ...*

Пирсон (1900/1956, с. 353) сопроводил вывод своего знаменитого критерия соответствия хи-квадрат резкими комментариями по поводу производимой обработки астрономических и геодезических наблюдений и стрельбы по мишеням. Упомянув “современные учебники по теории ошибок“, он заявил, что нормальный закон обычно выводится аналитически и что авторы

*Как правило приводят* [лишь] *скудные данные о его соответствии с действительными измерениями ... Возможно больше всех здесь провинился покойный Сэр Дж. Б. Эйри. ...*

**6.5. Видоизменения нормального закона.** По крайней мере некоторые из тех, которые видоизменяли нормальный закон, ошибочно полагали, что предлагают всеобъемлющий закон ошибок (§ 6.5.3). Далее, интересно, что только один автор (§ 6.5.6) предложил плотность в виде ряда Грама – Шарлье типа А, но и он впоследствии молчаливо отказался от своей рекомендации. При некоторых общих условиях этот ряд описывает распределения, близкие к нормальному.

**6.5.1.** Курно (1843, § 132) впервые рассмотрел ряд (астрономических или геодезических?) наблюдений с различными мерами точности, вначале, в § 81, исследовав подходящую урновую задачу. Пусть $n_1$, $n_2$, … наблюдений обладают плотностями $f_1(x)$, $f_2(x)$, … Тогда для всего ряда плотность будет равна

$$f(x) = [n_1 f_1(x) + n_2 f_2(x) + \ldots]/(n_1 + n_2 + \ldots].$$

Курно не указал, состояло ли различие этих плотностей только в параметрах точности, да и тип этих плотностей также остался неизвестным, хотя в §130 он привел рисунок (но не формулу) кривой, напоминающей нормальную и описывающую разброс ошибок наблюдений.

Заметим, что $f(x)$ соответствует смеси выборок, т. е. смеси значений некоторых случайных величин, а не их сумме. Именно сумма нормальных случайных величин с различными параметрами, но не их смесь снова нормальна (с определенными параметрами), см. § 6.5.5.

**6.5.2.** Де Морган (1864) попытался обобщить нормальный закон. На с. 410 он предположил, что закон ошибок может быть записан в виде

$$y = \varphi(x) = \sqrt{c/\pi}\ (p + qx^2 + rx^4 + \ldots) \exp(-cx^2). \qquad (4)$$

Введя моменты $A_0$ ( = 1), $A_2$, $A_4$, … он выписал для них соответствующие уравнения с неизвестными $c, p, q$, … Если (как он решил) требуется определить лишь два неизвестных, их можно будет установить по $c$ и моментам $A_0$ и $A_2$. Подставив полученные оценки в следующее уравнение, которое соответствовало $A_4$, Де Морган вывел квадратное уравнение с неизвестным $c$, заметил, что его корни действительны при

$$3A_2^2 - A_4 \geq 0 \qquad (5)$$

и выбрал соответствующее $c$.

Соотношение (5) означало, что эксцесс плотности (4) оказался либо отрицательным, либо нулем. Последующая практика, а именно более частое появление крупных ошибок, чем соответствовало бы нормальному закону (§ 6.4.2с), подсказала, что закон ошибок должен иметь положительный эксцесс, однако Де Морган (с. 420) предположил противное[10] и его предложение никогда непосредственно не применялось. Никто даже не вспомнил его, хотя именно он первым обобщил нормальный закон.

**6.5.3.** С 1873 по 1887 гг. несколько авторов, не упоминая Курно (§6.5.1) [Пирс; Стон, *Monthly Notices Roy. Astron. Soc.*, 1873 – 1874; Глейшер, там же, 1874; Эджворт, *Phil. Mag*., 1883 и 1887; Ньюком)[11], заявляли, что наблюдения некоторого данного ряда могут подчиняться нормальным законам с различными мерами точности, см. Хартер (1977). Мы опишем лишь попытку Ньюкома (1886, с. 351), который принял “весьма вероятное предположение”, т. е. что закон ошибок был “смесью [нормально распределенных] наблюдений” с различными мерами точности $h_i$, которые появляются с вероятностями $p_i$:

$$\psi(x) = (1/\sqrt{\pi})\ [p_1h_1 \exp(-h_1^2x^2) + p_2h_2 \exp(-h_2^2x^2) + \ldots]. \qquad (6)$$

Параметр нормального закона таким образом стал дискретной случайной величиной, однако величины $h_i$ и $p_i$, равно как и $n$, могли быть установлены лишь субъективно.

Ньюком далее указал, что неизвестную “величину” $\alpha$ следует определять по наблюдениям $x_1, x_2, ..., x_m$ в соответствии с условием

$$\int_{-\infty}^{\infty} (x-\alpha)^2\psi(x-x_1)\ \psi(x-x_2) \ldots \psi(x-x_m)dx = \min$$

и ввел упрощающие предположения, которые означали (Хюлм и Симмс 1939, с. 644), что оценка константы $\alpha$ соответствовала условию

$$\psi(\alpha - x_1)\ \psi(\alpha - x_2)\ \dots\ \psi(\alpha - x_m) = \max.$$

Ясно, что Ньюком ввел функцию потерь; по его собственным словам (1886, с. 348 прим.) он следовал Гауссу**[12]**.

Пирсон (1894), не упоминая Ньюкома, исследовал родственную задачу, – разложение анормальных кривых на нормальные составляющие. Он (с. 74) доказал, что “кривая, которая раскладывается на две нормальные составляющие, может быть разложена одним и только одним способом”, – для чего, однако, требуется решить алгебраическое уравнение девятой степени.

**6.5.4.** Леман-Филе (1887) видоизменил рассуждение Ньюкома, предположив, что $h$ – непрерывная (и несомненно положительная) случайная величина со своей собственной нормальной плотностью $\varphi_2(x)$, так что ошибки наблюдений оказались распределенными по закону

$$\varphi_1(x) = (h/\sqrt{\pi}) \exp(-h^2x^2),\ \varphi_2(x) = (k/\sqrt{\pi}) \exp[-k^2(h - h_0)^2]$$

и стало быть

$$P(v \le x \le v + dv) = cdv \int_{-\infty}^{\infty} \varphi_1(v)\ \varphi_2(h)\ dh.$$

Огородников (§ 6.5.6) улучшил эту рекомендацию, но не упомянул ее автора.

**6.5.5.** Эддингтон (1933, с. 277) весьма просто доказал, что эксцесс распределения (6) Ньюкома положителен; это в частности означало, что оно не было нормальным. Идельсон (1947, с. 309) назвал теорему Эддингтона и ее обобщение (Огородников) “одним из важнейших ... результатов современной теории ошибок”.

**6.5.6.** Огородников (1928; 1929a) предположил, что законом ошибок является

$$u(x) = (1/\sqrt{\pi}) \int_0^{\infty} hf(h) \exp(-h^2x^2)\ dh,$$

где $f(h)$ – не обязательно нормальная плотность распределения $h$. Далее, сославшись на Шарлье, он принял, что

$$u(x) = \varphi + \beta_3\sigma^3\varphi^{(3)} + \beta_4\sigma^4\varphi^{(4)} + \dots, \qquad (7)$$

где $\varphi$ – нормальный закон $N(0; \sigma)$, $\varphi^{(i)}$ – его $i$-я производная и $\sigma$ – средняя квадратическая ошибка наблюдений.

Он заметил, что нечетные коэффициенты $\beta_i$ ($i = 3, 5, \dots$) равны нулю и доказал, что $\beta_4$ неотрицательно. Вот его рассуждения: по своему знаку $\beta_4$ совпадает с $(\nu_4 - 3\nu_2^2)$,

$$\nu_2 = \int_{-\infty}^{\infty} x^2u(x)dx = (1/2)\int_0^{\infty} [f(h)/h^2]\,dh, \quad (8)$$

$$\nu_4 = \int_{-\infty}^{\infty} x^4u(x)dx = (3/4)\int_0^{\infty} [f(h)/h^4]\,dh, \quad (9)$$

$$\nu_2 = (1/2)\int_0^{\infty} \sqrt{f(h)}\,[\sqrt{f(h)}\,/h^2]\,dh,$$

$$\nu_2^2 \le (1/4)\int_0^{\infty} f(h)\,dh \int_0^{\infty} [f(h)/h^4]\,dh = (1/4)\int_0^{\infty} [f(h)/h^4]\,dh$$

и $(\nu_4 - 3\nu_2^2) > 0$, – равняется нулю только для нормального закона $u(x)$, при котором $h$ уже не является случайной величиной.

Огородников (1928, с. 16) заметил, что кривые плотности для скоростей собственных движений звезд "обладают весьма выраженным положительным эксцессом" и объяснил это тем, что средние скорости звезд различных спектральных классов весьма различны, ср. § 6.4.1b. Он добавил, что ряды наблюдений с отрицательным эксцессом вообще не известны. Интересно, впрочем, что ряды геодезических измерений иногда всё-таки обладают отрицательным эксцессом, видимо ввиду установленного инструкциями и соблюдаемого требования отбраковки уклоняющихся результатов (Кемниц 1957).

Наконец, Огородников предложил $\mathrm{E}h^2$ в качестве апостериорного веса наблюдений:

$$\mathrm{E}h^2 = \int_0^{\infty} h^3f(h)\,dh \exp(-h^2x^2)\,dh \div \int_0^{\infty} hf(h)\,dh \exp(-h^2x^2)\,dh =$$
$$u'(x)/[2xu(x)].$$

.

Это четная положительная величина, убывающая при $|x| \to \infty$. В позднейшей статье Огородников (1929b) обобщил свое исследование, предположив, что

$$u(x) = (1/\sqrt{\pi})\int_0^{\infty} h\,dh \int_{-\infty}^{\infty} f(h;a) \exp[-h^2(x-a)^2]\,da.$$

Разложений вида (7) он больше не стал использовать и ввел некоторые ограничения. К сожалению, на разложение анормальных кривых (Пирсон, § 6.5.3) он не сослался. Возможно, что в то время астрономы не были знакомы с работами английских статистиков.

### 6.6. Теория ошибок и статистика

**6.6.1.** Истинное значение и среднее состояние. Обойти понятие истинного значения при измерении констант вряд ли возможно, статистики же его почти не используют. В метрологии Эйзенхарт (1963/1969, с. 31), который предпочел не укоренившийся термин *целевое значение*, описал его как "предельное среднее в представляемом идеальном процессе". Это определение восходит к Фурье (1826/1890, с. 534), который заявил, что "истинный объект изучения" это предел среднего арифметического. В качестве следствия оказалось, что истинное значение константы непременно включает неизбежные остаточные систематические ошибки. Действительно (Эйзенхарт):

*Масса стандарта массы ... по определению ... это масса его металлического вещества плюс масса среднего объема воздуха, адсорбированного на его поверхности при стандартных условиях*.

Аналогично, угол фигуры триангуляции можно определить как соответствующий "теоретический" угол, искаженный метеорологическими условиями, обычными для данного района.

Уже в XVIII в. несколько авторов высказали мысль о том, что среднее арифметическое стремится к соответствующему истинному значению и тем самым косвенно ввели позднейшее определение Фурье. Так, Ламберт (1765b, § 3), ср. §.2.4.4:

*Допустимо принять, что каждый опыт может быть так же легко ошибочен в одну и в другую сторону и что равновозможны равные по величине уклонения по обе из них. Если это предположить, то легко доказать, что среднее из многих опытов должно приблизиться тем ближе к истинной величине, чем больше опытов повторяется. Ибо из всех случаев, которые можно здесь представить себе, наиболее возможен тот, при котором равные по величине уклонения происходят одинаково часто по обе стороны*.

Лаплас (1795, опубл. 1812/1912, с. 161) сформулировал аналогичную идею (и повторил ее в последующих сочинениях, хотя и не в *Опыте философии*):

*При неопределенном* [неограниченном] *увеличении числа наблюдений или испытаний их средний результат сходится к определенному члену* [числу], *так что, взяв сколь угодно малый интервал в одну и в другую сторону от него, вероятность, что средний результат окажется в нем, будет в конце-концов отличаться от достоверности меньше, чем на любую заданную величину. Этот член и есть сама истина, коль скоро положительные и отрицательные ошибки совершаются с одинаковой легкостью*.

Понятие истинного значения применяется и в геофизике (геомагнетизм, ускорение силы тяжести), и в физике (скорость света в пустоте, масса электрона). Интересно также, что начиная с Гаусса (мера точности наблюдений, см. 1816, §§ 3 и 4) это же понятие употреблялось даже в таких случаях, когда ему непосредственно ничего не соответствовало в природе.

Определение Фурье эвристически напоминает подход Мизеса к понятию вероятности, и следует указать, что Мизес (1919/1964, с. 46), хоть и не интересовался теорией ошибок, в основном повторил Фурье:

*Истинное среднее значение это лишь величина, которая, в соответствии с понятием о вероятности, должна произойти в качестве среднего арифметического когда ряд извлечений* [из урны] *продолжается до бесконечности*.

Напомним, что именно в этой работе Мизес ввел свое статистическое определение вероятности.

В статистике стремление среднего к соответствующему теоретическому параметру называется предельным свойством состоятельности, и оно имеет место для линейных оценок вообще. Тем не менее, в нашем контексте это вряд ли имеет значение, пока и поскольку (как явно указал, например, Ламберт, см. выше) практика имеет дело с наблюдениями, ощибки которых обладают четной плотностью.

Статистика более абстрактна, чем теория ошибок, так как занимается исследованием средних условий или состояний (и закономерностей уклонений от них), ср. § 6.4.1с. Так, в 1817 г. Гумбольдт (по примеру Галлея) ввел изотермы и тем самым выделил климатологию из метеорологии[13]. Это, кстати, было блестящим применением предварительного исследования данных (§ 0.2).

Статистика обязана изучать и столь смутные или фиктивные средние как количество ежегодных рождений в данном государстве[14] или среднюю цену хлеба и некоторые авторы (Кетле 1846, с. 63 и 65) подчеркивали отличие между действительными понятиями и подобными величинами.

Длительное время признавалось существование особой дисциплины, *теории средних*, которая как раз и изучала и те, и другие средние (Шейнин 1986, с. 311 – 312). Ввел ее Кондорсе (1805/1986, с. 604), но так и не объяснил своей мысли достаточно ясно. Уже Кетле (см. выше) фактически присоединил теорию средних к статистике, а на самом деле она оказалась поделенной между статистикой и теорией ошибок. Одним из последних о теории средних вспомнил Гильберт (1901, § 6):

*Что же касается аксиом теории вероятностей, то мне казалось бы желательным, чтобы параллельно с логическим обоснованием этой теории шло рука об руку строгое и удовлетворительное развитие метода средних значений в математической физике, в частности в кинетической теории газов*.

Известно, что в своем развитии математика неизменно удалялась всё дальше от реальности и тем самым становилась всё полезнее для естествознания, а в последние десятилетия – и для гуманитарных наук. Ясно поэтому, что переход математической статистики от истинных значений к оценке параметров функций явился шагом вперед. Но подчеркнем, что понятие истинного значения, равно как и теория ошибок в целом остаются необходимыми для многих естественных наук.

**6.6.2.** Кривые плотностей. В конце XIX в., когда статистики начали вводить различные кривые для описания эмпирических распределений, им пришлось выводить их параметры, а не определять средние состояния (тем менее, устанавливать истинные значения, § 6.6.1). И вообще статистикам пришлось придумать свою собственную терминологию (выборочное среднее, дисперсия, стандартное отклонение), но вот от гауссовых изящных обозначений типа [*ab*] они отказались, напрасно тем самым затруднив себе использование классических формул МНКв.

Фишер (1922, с. 309 – 310) усилил описываемый процесс, введя состоятельные, эффективные и достаточные оценки и обвинил статистиков (мог бы добавить и геодезистов) в том, что они не различали выборочных параметров от теоретических, – и назвал последние истинными (с. 311)! Он таким образом, как, впрочем, и Гаусс (§§ 6.6.1), применил термин *истинное значение* к объектам, которые не существовали в реальном мире.

**6.6.3.** Теория корреляции. Ее развитие привело к дальнейшему отчуждению между теорией ошибок и статистикой. По Гауссу (§ 5.3.4), зависимость между двумя [линейными] функциями обусловлена существованием (частично) общих аргументов, – ошибок наблюдения. Он (1826, с. 147) также упомянул

> *Взаимную зависимость наблюденных величин ... вызванную ... условными* [т. е. исходными] *уравнениями, которым ... должны удовлетворять их истинные значения.*

Это, видимо, означало, что непосредственно измеренные величины могли в идеале быть независимыми, но вот их уравненные значения – никогда. Точка зрения статистиков была схожей. В 1888 г. Гальтон (Пирсон 1920/1970, с. 199) утверждал, правда, в биологическом контексте, что

> *Корреляция должна быть следствием того, что вариации двух органов были частично обусловлены общими причинами.*

Впрочем, современная статистика больше интересуется соотношением причин и следствий и в вопросе о зависимости отошла от теории ошибок. Пирсон (с. 187) там же заявил, что

> *Если исключить условные уравнения, то в работах Гаусса нет и следа внутренней связи между наблюденными физическими величинами, а именно это и есть основополагающее понятие статистики.*

В то же время (Эйзенхарт 1978, с. 382) “разработанный Гауссом математический арсенал ... можно было непосредственно использовать в корреляционном анализе”. *Можно* было!

Разочаровавшись в разрабатываемой теории корреляции, Каптейн (1912) попытался ввести свой вариант корреляции в астрономию, а именно установить количественную меру связи между двумя функциями, зависящими от частично совпадающих аргументов. Его статья не была замечена, быть может потому, что он не сослался на Гаусса, видимо не представляя себе, что придал его взглядам количественное выражение. Тем не менее, выводы Каптейна без всякого упоминания о нем используются в геодезии, например для оценки степени зависимости двух смежных цепей триангуляции с общими базисом и астрономическим азимутом.

**6.6.4.** Разрыв между теорией ошибок и статистикой. Он вряд ли преодолим. Кроме указанных выше обстоятельств, разрыв обусловлен еще и непараметрическим подходом к погрешностям наблюдения, который принят в теории ошибок со времени второго гауссова обоснования МНКв и который соответствует реальному положению.

Для прикладников полностью неудачной оказалась книга Линника (1958). Он изложил теорию ошибок на математико-статистическом уровне, но не озаботился хоть как-то сохранить старое, и получается так, что существуют две теории ошибок, – классическая и математико-статистическая.

**Примечания**

**1.** *Этот случай отличается от схемы обычного уравнивания не принципиально, а скорее формально, и действительно … легко может быть приведен к* [ней]. Гаусс (1826, с. 147).

**2.** Бессель (§ 3) заметил, что случайные переменные должны быть независимы, но не упомянул это ограничение непосредственно при доказательстве своей теоремы. Несколько авторов подтвердили его видимо забытый или даже незамеченный результат, а Зейдель (1863, с. 326) сформулировал его без доказательства и привел ошибочную формулу для меры точности суммы двух нормальных законов.

Чубер (1903, с. 23), который сослался на Пиццетти и Линделёфа, заявил (и доказал), что

*Если независимые ошибки наблюдения … по-отдельности следуют нормальному закону, то их однородная линейная функция подчиняется … закону того же вида …*

Он назвал это положение “одной из основных теорем” теории ошибок. Сампсон (1913, с. 170) доказал ту же теорему и назвал двух своих предшественников. Он придал некоторое значение воспроизводству формы, но заметил, что трудно уточнить, что означает выражение *та же форма*.

**3.** Обозначим оценки одной и той же величины, полученные по МНКв и какому-то иному способу через $\alpha_1$ и $\alpha_2$. Тогда эффективность второй оценки будет равна $D\alpha_2/D\alpha_1 \le 1$, ср. конец § 5.3.5.

**4.** В конце 1940-х годов в *Трудах* ЦНИИ геодезии, аэросъемки и картографии появились отчеты об исследовании действия МНКв. Уравненные углы звена триангуляции искажались, после чего звено уравнивалось заново. Искажающие поправки выбирались по числу углов в соответствии с нормальным распределением с подходящими параметрами и распределялись случайным образом. Иначе говоря, использовался метод Монте Карло (или статистических испытаний), который, однако, авторы называли *методом искаженной модели*. Насколько нам помнится, результаты показали надежность МНКв, хотя, конечно же, возврата к прежней *модели* не могло произойти.

**5.** Несколько раньше Курно (1843, § 136 прим.) предложил применить “эту теорию к определению атомных весов или химических эквивалентов …” и косвенно уточнил, что имел в виду теорию ошибок.

**6.** Ср. также Милль (1843, с. 490):

*Весьма небольшое улучшение исходных данных при помощи лучших наблюдений или более полного учета специальных обстоятельств более полезно, чем самое тщательное применение исчисления вероятностей, основанное на данных в их прежнем, менее качественном виде.*

Аналогичное мнение высказал еще Лейбниц в письме 1703 г. Якобу Бернулли (Джини 1946, с. 405), но не следует всё же противопоставлять обстоятельства и вычисления, тем более, что первые могут вначале оставаться недостаточно известными.

**7.** В данном случае отбрасывание четырех экспериментов из пяти не имело значения, но интересно, что Джоуль не посчитал нужным обосновать свое решение. Ср. позднейшее утверждение Менделеева (1895, с. 159):

*Когда же одно из чисел представляет больше гарантий точности, чем другие, оно одно должно быть взято, оставляя безо всякого внимания числа, заведомо представляющие или худшие условия опыта и наблюдения, или какие-либо поводы к сомнению. … Брать во внимание с тем или иным “весом” худшие числа – значит … нарочно аортить лучшие из чисел …*

Но далеко не всегда можно определить, какие именно результаты хуже других.

**8.** Не введя никакой меры корреляции, Зейдель (1865; 1866) первым количественно исследовал зависимость между несколькими факторами. Именно, он определял, зависит ли заболеваемость тифом от уровня грунтовых вод, а затем от этого же уровня и, дополнительно, от количества осадков. На его работу обратили внимание лишь врачи (Вейлинг 1975; Шейнин 1982, §§ 7.4.2 – 7.4.3).

В то время статистики качественно исследовали наличие или отсутствие зависимостей по характеру изменения (монотонному или нет) интересующей их функции при монотонном изменении ее аргумента. Примеры: вероятность осуждения обвиняемого в зависимости от его социального положения, образования и т. д. (Кетле 1836, том 2, с. 213); возрастание смертности в больницах с количеством коек в них, т. е. с ухудшением санитарных условий (Дж. Симпсон 1869 – 1870, с. 399).

**9.** Напомним (§ 3.6.1), что Лаплас был готов признать нормальный закон в качестве закона ошибок.

**10.** См. также мнение Огородникова (§ 6.5.6). Термин *эксцесс* ввел Пирсон (1894, с. 93), определивший его как $(A^4 - 3A_2^2)/3A_2^2$. Там же он ввел словообразования *стандартное уклонение* и *нормальная кривая* (с. 75 и 72), см. наш § 4.6.2.

Де Морган заметил, что $p > 0$ и $q < 0$. Для больших значений $|x|$ его функция (4) оказывалась таким образом отрицательной, но он (с. 421) заявил, что "количественное последствие этого слишком незначительно и не требует внимания" и что он вообще "еще не встречал ни единой задачи", в которой "имело бы смысл" истолковывать отрицательную вероятность.

Без всяких пояснений он добавил совсем уже нелепое утверждение: некоторое событие, как он заметил, имело вероятность 2.5,

> *Что означало, что … оно должно произойти дважды и притом иметь равные шансы появиться или нет в третий раз.*

**11.** Несколько слов о мемуаре Эджворта (1883). Во-первых, он применяет особый термин *кривая вероятностей* для нормальной плотности, но в общем случае называет плотности *кривыми возможностей*. Во-вторых (с. 361), статистическая оценка для него является *ущербом* (evil). В узком и естественном смысле этот термин позднее применил Ньюком (1886). В третьих, Эджворт (с. 363) относится к сочетанию наблюдений как к определению наибольшей полезности. В четвертых, он (там же) каким-то образом полагает, что для каждого положительного значения аргумента $x$ и заданных плотностей $\varphi_1(x)$ и $\varphi_2(x)$

$$\int_0^x \varphi_1(x)dx > \int_0^x \varphi_2(x)dx.$$

**12.** Через год Хюлм (1940) опубликовал статью о проникновении математико-статистических идей в теорию ошибок.

**13.** Вообще же примерно между 1815 и 1915 гг. возникло не менее семи научных дисциплин, непосредственно связанных со статистикой: география растений, общая гигиена (предшественница экологии), звездная статистика, эпидемиология, зоогеография, кинетическая теория газов, биометрия, – и, по существу, теория ошибок.

**14.** Вот, действительно, несколько неизбежно возникающих вопросов: включать ли иностранцев, проживающих в государстве и/или своих граждан, находящихся за рубежом? Как учесть различие между данными регистрации и фактическим количеством рождений?

**Основная литература**

Бессель (1961), Ку (1969), Ньюком (1886), Шейнин (1984a; 1984b; 1994b)

**7. Гельмерт**

Трактат Гельмерта (1872) – лучший источник для изучения состояния классической теории ошибок и МНКв. К моменту выхода в свет его второго издания (1907 г.; третье, посмертное, появилось в 1924 г.) интересы Гельмерта изменились, он начал изучать фигуру

Земли и не попытался использовать зарождавшуюся в то время теорию корреляции. Мы рассмотрим его основные результаты, которые он сообщал в своих многочисленных статьях; в меньшей степени они отражены во втором издании его трактата.

**7.1. Отбраковка улоняющихся наблюдений.** Иордан (1877) принял четный трехчлен

$$\psi(x) = a + bx^2 + cx^4, |x| \le M$$

в качестве приближения нормального распределения. Полагая, что $\psi'(M) = 0$, он вычислил все три его параметра в функции $M$ и заметил, что

$$m^2 = 2\int_0^M x^2\psi(x)dx = (1/7)M^2 \approx M^2/2.65^2, M \approx 2.65m.$$

Соответственно, он рекомендовал отбраковывать все наблюдения, уклоняющиеся от своего среднего более, чем на $3m$. Так было введено знаменитое правило трех сигма[1].

Гельмерт (1877) существенно возразил. Он заметил, что приближение к нормальному закону окажется тем лучше, чем больше членов будет включено в $\psi(x)$ и тем бóльшим станет отношение $M/m$. Далее, он указал, что не подобный трехчлен, а сочетание нескольких независимых элементарных ошибок, равномерно распределенных на одном и том же отрезке, является естественным источником появления нормального закона, ср. прим. 18 гл. 5. В третьих, Гельмерт (с. 143) разумно подчеркнул, что $M$ зависит от числа наблюдений $n$ и потому является ожидаемой, а не установленной заранее величиной. И все же Гельмерт согласился с новым правилом, если только $n = 10 – 100$ и тем самым пошел несколько дальше, чем Гаусс (§ 5.4.3).

**7.2. Выявление систематических ошибок**. Начиная, видимо, с Тихо Браге, астрономы неизменно обращали внимание на систематические ошибки, – на их исключение и в процессе наблюдений, и при обработке результатов. Рассматривая последнюю задачу, Лаплас (§ 3.5.2) рекомендовал выделять среднее значение систематической ошибки, а Гельмерт (1872, с. 257; 1875с, с. 147 и 151), не ссылаясь ни на кого, продолжил изучение этой темы.

Впрочем, его основной результат появился в 1905 г., когда он предложил несколько правил.

**7.2.1.** Пусть ошибки наблюдений равны $v_i\varepsilon_i$, $v_i$ и $v_i = 1$ или $– 1$ и притом $f$ раз $v_iv_{i+1} = 1$, $i = 1, 2, \ldots, n – 1$ и $w$ раз это же произведение равно $– 1$. Тогда

$$f – w = v_1v_2 + v_2v_3 + \ldots + v_{n–1}v_n, \text{E}(f – w) = 0, \text{E}(f – w)^2 = n – 1$$

и частости значений $(f – w)$ равны удвоенным числам соответствующих биномиальных коэффициентов. Так, при $n = 5$, равенства $f – w = – 4, – 2, 0, 2, 4$ имеют место 2, 8, 12, 8 и 2 раза. Гельмерт не привел нетрудного доказательства этого. Он (с. 603) далее вывел нормальное приближение для плотности распределения указанной разности, заметив, что этот результат принадлежал Зеелигеру (1900). Как и в случае других предложенных им правилах, он исходил из того, что для случайной величины $\xi$ и ее математического ожидания $E\xi$ неравенство $P(|\xi| \geq \sigma_\xi)$, где $\sigma_\xi$ – ее средняя квадратическая ошибка, имеет место с вероятностью примерно равной 0.68, и что если оно не выполняется, то можно подозревать наличие систематических ошибок. Он таким образом молчаливо исходил из нормально распределенных величин.

Не обсуждая нескольких других предложенных им правил, перейдем к критерию Аббе и его видоизменению, рекомендованному Гельмертом.

**7.3. Продолжение: критерий Аббе.** Пусть ошибки наблюдения будут $\varepsilon_i$, $i = 1, 2, \ldots, n$ и, далее, обозначим

$$\Delta = [\varepsilon\varepsilon],\ \theta = (\varepsilon_1 - \varepsilon_2)^2 + (\varepsilon_2 - \varepsilon_3)^2 + \ldots + (\varepsilon_{n-1} - \varepsilon_n)^2 + (\varepsilon_n - \varepsilon_1)^2,$$
$$\mu = \theta/\Delta.$$

Аббе (1863, с. 80 – 81) заметил, что $E\mu = 2$ и что чем слабее было влияние систематических ошибок, тем меньше окажется $|\mu - E\mu|$. Действительно, эта разность не равна нулю при $E(\varepsilon_i - \varepsilon_{i+1}) \neq 0$, т. е. когда $E\varepsilon_i \neq 0$.

По существу систематические ошибки влияли только на первую из двух величин, $\Delta$ и $\theta$. Поэтому, как указал Аббе (с. 59 – 60), его статистика $\mu$ была чувствительна к “односторонним” ошибкам, т. е. к “причинам, действующим закономерно”.

Гельмерт (1905, с. 606 – 612) явно исходил из критерия Аббе, но не сослался на него.

**7.3.1** Иная статистика (с. 606). Пусть теперь $A$ и $B$ обозначают то, что раньше обозначалось через $\Delta$ и $\theta$ и $C = A - B/2$, тогда, если $m$ – средняя квадратическая ошибка наблюдения,

$$EA = nm^2,\ EC = 0,\ EC^2 = m^4 n.$$

В предположении, что

$$A = EA \qquad (1)$$

Гельмерт получил

$$nm^2 - B/2 = \pm\, m^2\sqrt{n},\ B = 2m^2 n\,[1 \pm (1/\sqrt{n})],\ A = (B/2)[1 \pm (1/\sqrt{n})]^2,$$

$$E[\sqrt{A/n} - \sqrt{B/2n}\,] = 0,\ E[\sqrt{A/n} - \sqrt{B/2n}\,]^2 = m^2/4n. \qquad (2)$$

Условие (1) было необходимо, потому что $A$ оставалось неизвестным, тогда как $B$ можно было вычислить, поскольку $(\varepsilon_i - \varepsilon_j) = (v_i - v_j)$, т. е. разность истинных ошибок равна разности соответствующих остаточных свободных членов исходных уравнений. Формулы (2) были, конечно, приближенными ввиду допущения (1).

**7.3.2.** Другое видоизменение статистики (с. 607). Гельмерт теперь ввел величину $B$, равную прежней, но без последнего слагаемого и получил

$$\mathrm{E}\,[\sqrt{A/(n-1)} - \sqrt{B/[2(n-1)]}\,]^2 = m^2/[4(n-1)].$$

Разность $C = A - B/2$ включает величины $\varepsilon_i\,\varepsilon_{i+1}$, которые, как и в § 7.3.1, связаны лишь со своими смежными членами, притом тем слабее, чем больше $n$. Указав на это, Гельмерт (с. 608) дополнительно привел без доказательства формулу

$$\mathrm{E}C^4/(\mathrm{E}C^2)^2 = 3 + k/(n-1),\ k = \mathrm{Const},$$

так что при $n \to \infty$ распределение $C$, как оказалось, стремилось к нормальному.

Это непонятно, потому что только что выписанное уравнение не могло быть получено без предположения о каком-либо соотношении между $\mathrm{E}C^4$ и $\mathrm{E}C^2$. Тем не менее, для нормально распределенных ошибок $\varepsilon_i$ утверждение Гельмерта представляется справедливым. Так (Нейман и др. 1941, с. 155), $B/(n-1)$ “видимо приближается к нормальному распределению”, а первое слагаемое в $C$, т. е. $[\varepsilon\varepsilon]$, наверняка обладает этим свойством.

**7.3.3.** Гельмерт (с. 608 – 610) рассматривает не ошибки, а остаточные свободные члены. МНКв требует выполнения условия $[vv]$ = min, и он (с. 608) посчитал, что возникающее при этом “стеснение” влияет систематически. Но исследование этого обстоятельства оказалось трудным и он ограничился случаем непосредственных наблюдений, притом лишь для нормального распределения, и применил ту же статистику $B$, что и в § 7.3.2.

**7.4. Суммы натуральных степеней ошибок.** Гельмерт (1875а; 1875b)[3] определял распределение суммы

$$\varepsilon_1^{\ m} + \varepsilon_2^{\ m} + \ldots + \varepsilon_n^{\ m}$$

ошибок наблюдений $\varepsilon_i$, распределенных нормально или равномерно, и, кроме того, произвольно, но при $n \to \infty$. В своей второй, основной статье он рассмотрел

**7.4.1.** Произвольное распределение и конечное $n$. Гельмерт начал со случая $n = 1$ и вывел распределение величины $\varepsilon_1^{\ m}$. Он, конечно же, не применил ныне стандартной формулы для вычисления плотности функции случайной величины.

Случай $n = 2$ был уже утомительным. Затем, переходя к произвольному $n$, Гельмерт выписал соответствующий кратный интеграл и вычислил его при помощи разрывного множителя Дирихле.

**7.4.2.** Равномерное распределение; $n = 1$ и 2 и $m = 1$, 2 и 3. Случай $n = 2$ снова оказался трудным.

**7.4.3.** Нормальное распределение; $n = 1$ и 2 и $m = 1$, 2 и 3.

**7.4.4.** Нормальное распределение; $m = 2$ и конечное $n$. (с. 202 – 205). Исходя из только что полученного результата (§ 7.4.3), Гельмерт по индукции вывел распределение хи-квадрат. Хальд (1960, с. 258 – 261) модернизировал этот вывод.

Пирсон (1931) специально указал, что Гельмерт опередил его в выводе указанного распределения и описал этот вывод. Он, однако, не упомянул основного в этом смысле мемуара (1876b) и лишь рассмотрел другую статью (1875a).

**7.4.5.** Предельная теорема для произвольного распределения (с. 205 – 211). Следуя за Пуассоном (1837, с. 267) и Глейшером (1872), рассмотревшим случай $m = 1$, Гельмерт вывел соответствующую предельную теорему, которая совпала с теоремой Гаусса (§ 5.2.2). Его вывод не был убедителен, поскольку он принял несколько предположений, но не определил вытекающих из них погрешностей.

**7.4.6.** Предельные теоремы для равномерного и нормального распределений (с. 211 – 213). Гельмерт соответствующим образом специализировал формулу Гаусса (§ 5.2.2).

**7.4.7.** Мера точности для нормального распределения (с. 214 – 215). В этом случае

$$S_m = \int_0^a \varepsilon^m \psi(\varepsilon) d\varepsilon = \Gamma[(m+1)/2]/h^m\sqrt{\pi}.$$

Здесь $\psi(\varepsilon) = \psi_1(\varepsilon) + \psi_2(\varepsilon)$ – сумма плотностей для отрицательных и положительных ошибок при $|\varepsilon_i| \le a$ $(a > 0)$. Меру $h$ Гельмерт оценил, приняв

$$S_m = (1/n)\sum|\varepsilon_i|^m.$$

**7.5. Распределение хи-квадрат.** Таким образом (§ 7.4.4), Гельмерт вывел распределение хи-квадрат и Пирсон признал его первенство. Более того, Гельмерт опубликовал свое открытие (без доказательства) в прежней работе (1875a), но всё же у него были предшественники.

**7.5.1.** Лаплас (см. формулу (3.24a) в § 3.6.2) вывел для параметра точности (Ланкастер 1966, с. 120)

*Распределение типа гамма, хотя и не обычное для суммы квадратов, а распределение точности в бейесовском предположении.*

Он мог бы легко установить распределение величины

$$2h[TT]/3 = [TT]/3\sigma^2$$

(обозначения см. § 3.6.2), получив $\chi^2_{n+2}$ вместо $\chi^2_{n-1}$ как при стандартном “не-бейесовском” подходе). Заметим также, что Лаплас, см. формулу (3.2.4b), предположил, что $n$ – большое число.

**7.5.2.** Следующим в 1852 г. был Бьенеме (Хейде и Сенета 1977, § 4.3), но его исследование не было связано с ошибками наблюдения.

**7.5.3.** Аббе (1863) вывел распределение хи-квадрат в теории ошибок (Шейнин 1966). Гельмерт не указал этого ни в 1876 (§ 7.4.4), ни в 1905 г. (§ 7.3). Кендалл (1971) описал работу Аббе современным языком.

Сколь ни существенно распределение хи-квадрат для статистики, в классической теории ошибок оно почти не применялось, см. § 5.4.2.

**7.6. Формула Петерса.** Для случая $n$ непосредственных наблюдений и нормального распределения Петерс (1856) вывел формулу средней абсолютной ошибки единичного веса

$$\theta = \sum_{i=1}^{n} |v_i| / \sqrt{n(n-1)}. \qquad (3)$$

Здесь $v_i$ – уклонения наблюдений от арифметического среднего.

Гельмерт (1875b) обосновал эту формулу заново, потому что Петерс молчаливо и ошибочно предположил, что величины $v_i$ взаимно независимы, см. ниже.

**7.6.1.** Непосредственные наблюдения, $n = 2$. Поскольку $|v_1| + |v_2| = |\varepsilon_1 - \varepsilon_2|$, где $\varepsilon_i$ – истинные ошибки, то Гельмерт вычислил

$$E|\varepsilon_1 - \varepsilon_2| = (h/\sqrt{\pi})^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[-h^2(\varepsilon_1^2 + \varepsilon_2^2)]\, |\varepsilon_1 - \varepsilon_2| d\varepsilon_1\, d\varepsilon_2 = \sqrt{2}/h\sqrt{\pi}.$$

В этом случае формула Петерса оказалась верной, так как

$$E|\varepsilon| = (2h/\sqrt{\pi}) \int_{-\infty}^{\infty} |x| \exp(-h^2x^2)dx = 1/h\sqrt{\pi}.$$

**7.6.2.** Непосредственные наблюдения, $n$ произвольно. Здесь

$$v_i = \varepsilon_i - \sum\varepsilon_i/n, \quad \sum|v_i| = \sum|\varepsilon_k - \sum\varepsilon_i/n|. \qquad (4a;\ 4b)$$

Гельмерт вычислил соответствующий кратный интеграл при помощи разрывного множителя Дирихле и снова получил формулу Петерса. Его непростые выкладки сводились к приравниванию $\sum|v_i|$ мнимой части интеграла

$$2/[\pi\, (h/\sqrt{\pi})^n] \int_{-\infty}^{\infty} du/u \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} d\varepsilon_1 d\varepsilon_2 \ldots d\varepsilon_n \cdot$$

$$\sum_{k=1}^{n} \{(\varepsilon_k - \sum \varepsilon_i/n) \exp[-h^2[\varepsilon\varepsilon] + iu\,(\varepsilon_k - \sum \varepsilon_i/n)]\}.$$

Гораздо проще, однако (Дейвид 1957, с. 27), вспомнить, что для нормального распределения

$$E\sum|v_i| = \sqrt{(n-1)/n}\,/(h/\sqrt{\pi})$$

и что ожидаемое значение суммы всех уклонений в $n$ раз больше. Это замечание основано на устойчивости нормального закона: поскольку $v_i$ являются линейными формами $\varepsilon_i$, см. формулу (4a), они также распределены нормально. Гельмерт нигде не применил этого свойства нормального закона и, видимо, не знал о нем. И, указав, что при выводе нормального закона Гаусс (§ 5.1.3) молчаливо заменил истинные ошибки уклонениями, Гельмерт (1872, с. 75), конечно же, упустил прекрасную возможность упомянуть об устойчивости этого закона.

Чубер (1891а, с. 108) косвенно подтверждает наше мнение. Быть может следуя за Гельмертом или Меррименом (§ 5.1.3), но не ссылаясь ни на кого, он указал на тот же недостаток в выводе Гаусса, но в подробности не вошел. Позднее он (1903) доказал, что линейная функция независимых и нормально распределенных ошибок также нормальна, см. прим. 2 к гл. 6. Своей прежней работы 1890 г., см. § 3.6.3, он не упомянул.

**7.6.3.** Косвенные наблюдения при $m$ неизвестных ($m > 1$). В этом трудном случае Гельмерт не вывел заново формулу Петерса, а точнее, не обобщил ее, но доказал, что она занижает среднюю абсолютную ошибку. Он также указал, что эта ошибка менее чем $\mathrm{E}\sum|v_i|/(n - m)$, однако его пояснения слишком кратки. Ясно, впрочем, что

$$\mathrm{E}(\sum|v_i|)^2 > \mathrm{E}[vv] = (n - m)/2h^2,$$

так что

$$\mathrm{E}\sum|v_i| > \sqrt{n-m}\,/(h/\sqrt{\pi}) = \sqrt{n-m}\;\mathrm{E}|\varepsilon|,\ \mathrm{E}|\varepsilon| < \mathrm{E}\sum|v_i|/\sqrt{n-m}\,.$$

**7.6.4.** Точность формулы Петерса. Ограничиваясь случаем $n$ прямых наблюдений, Гельмерт (1876а) вычислил дисперсию $\theta$. Его исследование было нелегким и необходимым с теоретической точки зрения, но вряд ли существенным для практики; действительно, формула Гаусса (5.24) сохраняет силу для любого распределения. Много позже Фишер (1920, с. 761) независимо повторил результат Гельмерта.

**7.7. Формула Гаусса.** Гаусс (§ 5.3.7) вывел формулу для квадрата средней квадратической ошибки наблюдения

$$m^2 = [vv]/(n - k), \tag{5}$$

где $n$ и $k$ – количества наблюдений и неизвестных. Он также установил границы для дисперсий $m^2$ и в общем случае, и, отдельно, для нормального распределения, см. формулы (5.25) и (5.26).

Не применяя новых методов, Гельмерт (1904) исправил ошибку Гаусса, получив при $\rho = \nu_4 - 3s^4 \geq 0$ и $\rho \leq 0$ соответственно

$$\frac{\nu_4 - s^4}{n-k} - \frac{\rho}{n-k}\cdot\frac{k}{n} < \mathrm{D}m^2 < \frac{\nu_4 - s^4}{n-k}, \tag{6a}$$

$$\frac{\nu_4 - s^4}{n-k} < \mathrm{D}m^2 < \frac{\nu_4 - s^4}{n-k} + \frac{k}{n}\cdot\frac{3s^4 - \nu^4}{n-k}. \tag{6b}$$

Переписывая эти выражения, мы учли наше собственное замечание из § 5.3.7 и обозначили $\text{E}m^2 = \text{E}\varepsilon_i^2$ через $s^2$.

Колмогоров и др. (1947) независимо получили тот же результат. Они и Мальцев (1947) показали, что в обоих приведенных выражениях знак строгого неравенства можно заменить на $\le$.

Вспомним (§ 5.3.7), что формула Гаусса (5) вынужденно включала не $\text{E}[vv]$, а $[vv]$, Гельмерт (1876a) же вознамерился вывести эту формулу без указанной замены, однако ему пришлось зато ограничиться нормальным распределением и, более того, случаем прямых наблюдений ($k = 1$). Он вычислил вероятность получения ошибок $\varepsilon_i$ в виде

$$P_1 = (h/\sqrt{\pi})^n \exp(-h^2[\varepsilon\varepsilon])\, d\varepsilon_1 d\varepsilon_2 \ldots d\varepsilon_n$$

и заметил, что она максимальна при $(1/2h^2) = m^2 = [\varepsilon\varepsilon]/n$. Затем Гельмерт ввел уклонения $v_i$:

$$v_i = \varepsilon_i - \beta,\ \beta = (1/n)\sum\varepsilon_i, \qquad (7a;\ 7b)$$

причем, разумеется, их сумма равнялась нулю. Таким образом,

$$P_1 = n\,(h/\sqrt{\pi})^n \exp[-h^2[vv] + n\beta^2)]\, dv_1 dv_2 \ldots dv_{n-1} d\beta \qquad (8)$$

и вероятность получить уклонения $v_i$ оказалась равной

$$P_2 = \sqrt{n}(h/\sqrt{\pi})^{n-1} \exp(-h^2[vv])\, dv_1 dv_2 \ldots dv_{n-1}, \qquad (9)$$

а ее максимальное значение соответствовало равенству

$$(1/h\sqrt{2}) = m = \sqrt{[vv]/(n-1)}.$$

**7.8. Предвосхищение теоремы Стюдента – Фишера.** Гельмерт не обратил особого внимания на распределение (8), где, в соответствии со своим определением (7b), β было ошибкой среднего арифметического. Но именно эта формула показывает, что при его условиях $[vv]$ (и, следовательно, выборочная дисперсия) и среднее арифметическое не зависимы друг от друга. Указав на это, Краскл (1946) упомянул несколько современных выводов этой формулы и предложил свое собственное индуктивное доказательство.

Гельмерт предвосхитил Стюдента на несколько десятилетий[4].

**7.9. Точность средней квадратической ошибки.** Границы (6) относятся к дисперсии, тогда как Гельмерт (1876a) дополнительно исследовал среднюю квадратическую ошибку *m*. Как обычно, он ограничился нормальным распределением, для которого даже формула Гаусса (5.24) не давала необходимого ответа.

Гельмерт заметил, что вероятность неравенств $\beta \le [vv] \le \beta + d\beta$ равна интегралу от выражения (9), взятому в указанных пределах. Затем он ввел новые переменные[5]

$$t_1 = \sqrt{2}\,[v_1 + (1/2)v_2 + (1/2)v_3 + (1/2)v_4 + \ldots + (1/2)v_{n-1}],$$

$$t_2 = \sqrt{3/2}\,[\quad v_2 + (1/3)v_3 + (1/3)v_4 + \ldots + (1/3)v_{n-1}],\ \ldots,$$
$$t_{n-1} = \qquad \sqrt{n/(n-1)}\, v_{n-1},$$

так что $[vv] = \sum t_i^2$ и якобиан его преобразования оказался равным $J = \sqrt{n}$. Соответственно, сославшись на свой вывод распределения хи-квадрат (§ 7.4.4), Гельмерт получил

$$P(\beta \le \sum t_i^2 \le \beta + d\beta) = \frac{h^{n-1}}{\Gamma[(n-1)/2]}\beta^{(n-3)/2}\exp(-h^2\beta)d\beta,$$

$$\mathrm{E}\sqrt{[vv]} = \frac{h^{n-1}}{\Gamma[(n-1)/2]}\int_0^\infty \sqrt{\beta}\beta^{(n-3)/2}\exp(-h^2\beta)d\beta = \frac{\Gamma(n/2)}{h\Gamma[(n-1)/2]}.$$

Наконец, поскольку

$$\mathrm{E}[vv]/(n-1) = m^2 = 1/2h^2,$$

$$\mathrm{E}\{m - \frac{[vv]}{\sqrt{n-1}}\}^2 = (1/h^2)\{1 - \frac{\sqrt{2}\Gamma(n/2)}{\Gamma[(n-1)/2]\sqrt{n-1}}\}.$$

**7.10. Необходима ли несмещенность?** Спротт (1978, с. 194) заметил, что при оценке параметров несмещенность вряд ли нужна, а иногда и невозможна, а также (что уже непосредственно относится к теории ошибок), что, несмотря на несмещенность выборочной дисперсии оценки (5), средняя квадратическая ошибка смещена. Сам Гаусс (1828, § 3) оценивал точность угловых измерений не дисперсией, а средней квадратической ошибкой, и именно так геодезисты поступали с тех самых пор. По этой причине исследование Гельмерта (§ 7.9) было важным, но о нем, видимо, забыли.

Первым критиком формулы Гаусса (5.24) оказался Бертран (1888d, с. 281 – 282), который предложил “лучшую” оценку точности. Молчаливо исходя из нормального распределения, он вычислил $\mathrm{D}m^2$, забыв при этом о существовании формулы (5.26), и доказал, что для его оценки эта же величина была меньше. Смещенности своей оценки он не отметил.

Чубер (1891b, с. 460) обсудил пример Бертрана с Гельмертом:

> *Оценку этого ошибочного заключения мне удалось случайно обсудить с проф. д-ром Гельмертом, который высказал мне свои сомнения по поводу вышеуказанной разработки* [Бертрана] *и вскоре привел также доводы в пользу своих собственных результатов.*
> *Если они верны, то возражения против теории Гаусса окажутся весомыми. Именно, Бертран* [как и сам Гаусс] *оценил неуверенность формулы* [Гаусса] *по абсолютной величине ее средней квадратической ошибки, вместо того, чтобы основываться, как*

*это должно быть, на относительной величине этой ошибки …*

Таким образом, несмещенность не смущала ни Гельмерта, ни Чубера. Сам Чубер доказал, опять же для нормального распределения, что $\mathrm{D}m^2/m^2$ минимально для оценки Гаусса. В этом случае, стало быть, формула Гаусса (5.24) обеспечивает не только наименьшую несмещенную оценку, но и наименьшую относительную дисперсию. Важно ли это? Да, по крайней мере для практики. Эддингтон (1933, с. 280), к примеру, также высказался в пользу относительной дисперсии.

В настоящее время смещение статистических оценок допускается, по крайней мере в некоторых границах, см. также выше. Было бы, видимо, разумно распространить этот подход на обработку наблюдений, тем более, что (Спротт, см. выше) для оценки точности применяется не дисперсия, а смещенная средняя квадратическая ошибка. Напомним, однако, что соответствующее исследование Гельмерта (§ 7.9) относилось лишь к нормальному распределению.

В частном случае и для смещенной выборочной оценки дисперсии (5), т. е. при $k = 0$ вместо 1, Крамер (1946/1948, с. 382) получил выражение в терминах центральных моментов

$$\mathrm{D}(m^2) = \frac{\mu_4 - \mu_2^2}{n} - \frac{2(\mu_4 - 2\mu_2^2)}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}.$$

**Примечания**

**1.** Фурье (1826/1890, с. 541 – 543) рекомендовал принять в качестве “границы наибольших ошибок” в √2 раз бóльшую величину. Наиболее распространенное после правила трех сигм предложил Шовене (1863/1960, том 2, с. 558 – 566). Ожидаемое число переменных $\xi_i$, удовлетворяющих условию $|\xi_i| > x$, равно

$$n[1 - \varphi(x/\sigma)],\ \varphi(x/\sigma) = (1/\sqrt{2\pi}\ ) \int_0^{x/\sigma} \exp(-z^2/2)dz,$$

где σ считается известным и $n$ – число наблюдений. По Шовене, наибольшее $|\xi_i|$ отбрасывается, если

$$n[1 - \varphi(\max|\xi_i|/\sigma)] > 1/2.$$

И вот несколько дополнительных строк из неопубликованной рукописи Л. Н. Большева (1922 – 1978), с которой он ознакомил нас:

$$P(\max|\xi_i| > x) \le nP(|\xi_1| > x),\ P(\max|\xi_i| > x) \ge$$

$$nP(\xi_1 > x) - [(n(n-1)/2]P(|\xi_1|, |\xi_2| > x) \ge nP(|\xi_1| > x) - n^2[P(|\xi_1| > x)]^2$$

и поэтому

$1/4 \leq P(\max|\xi_i| > x) \leq 1/2.$

Авторы XX в. (Диксон 1962; Краскл 1960/1969, с. 348) признавали, что не знают общего решения задачи об отбраковке наблюдений. Статистические правила очевидно могут помочь здесь, но ни одно из них не надежнее своих предпосылок, выполнение которых трудно проверить. Особо требуется решить, как поступать с наблюдениями, которые не могут принадлежать к той же совокупности, что и все остальные. Барнетт и Люис посвятили целую книгу (1978) этой теме и заключили (1978/1984, с. 360), что

*В конце концов, главной задачей при изучении уклоняющихся наблюдений остается та, которая встретилась самым первым исследователям нашего предмета: какое наблюдение считать уклоняющимся и что с ним делать*?

**2.** Гельмерт с самого начала (1872, с. 12) сопровождал среднюю квадратическую ошибку двойным знаком, и его решение было воспринято геодезистами. Ни Гаусс (1816, §§ 2 и 8), ни Бессель (1838b, §§ 5, 6 и 12) не были в этом смысле последовательными.

**3.** Мы упоминаем первую статью в начале § 7.5.

**4.** Бертран (1888b; 1888c) также близко подошел к теореме Стьюдента – Фишера. Вторую его статью в этой связи упомянули Хейде и Сенета (1977, с. 67 прим.). Бертран уделил много внимания теории ошибок. Будучи еще новичком, он перевел соответствующие работы Гаусса на французский язык (Гаусс 1855), а его собственные труды по теории ошибок (заметки в *C. R. Acad. Sci. Paris* и его знаменитый трактат) относятся к 1887 – 1888 гг. Перевод Гаусса назван на титульном листе авторизованным, однако сам Бертран (1855) пояснил, что Гаусс (умерший в том же году) так и не успел прислать ему свои замечания.

К сожалению, не владея современными ему знаниями в астрономии и геодезии (ср. §§ 5.4.1 – 5.4.2), он часто ошибался и, подытоживая свои критические замечания, даже заявил (1888d, с. 222), что "приложение исчисления вероятностей к изучению ошибок наблюдения основано на вымысле".

Одно замечание Бертрана (там же, с. 267) мы полагаем интересным. Он указал, что для небольших по абсолютной величине значений аргумента нормальный закон может быть приближен квадратным двучленом. Мы вернемся к Бертрану в § 7.10.

**5.** Мы не останавливаемся на преобразовании Гельмерта, как оно называется в статистике.

**Основная литература**

Гельмерт (1872), Шейнин (1995а)

## 8. Устойчивые законы (Леви)

Устойчивые законы начали изучаться в 1920е годы и Леви был сооснователем их теории. Он же оказался единственным автором, заявившим, что эти законы необходимы для создания новой теории ошибок (которую он совершенно не понял), см. также § 0.2. Его рассуждения по этому поводу, которые мы постарались по возможности систематизировать, в основном содержались в двух методически сырых трудах (1924; 1925).

**8.1. Случайные ошибки.** Леви заметил, что их средние равны нулю (1924, с. 51; 1925, с. 278), что они "независимы и очень малы" (1924, с. 50), или, по крайней мере, что они проявляются "как суммы" таких ошибок (1925, с. 278).

Он (1925, с. 70 – 71 и 278) дважды заявил, что случайные ошибки распределены нормально, но уточнил (с. 73): почти нормальны. А на с. 279 заключил:

> *В конце концов случайная ошибка подчиняется закону Гаусса тем точнее, чем более точно подтверждаются условия* [ЦПТ].

О других возможностях Леви не подозревал. Много позднее он (1970, с. 71) указал по поводу теории ошибок, что в 1919 г. "лишь смутно помнил, что случайные ошибки подчиняются закону Гаусса".

Впрочем, Леви в основном интересовался иными распределениями и видимо, обсуждал главным образом наблюдения, искаженные систематическими влияниями.

**8.2. Точность наблюдений.** Точность можно всесторонне описывать лишь соответствующим законом распределения (1924, с. 78 – 79; 1925, с. 75 – 76). Эту верную мысль трудно, однако, использовать. Те, кто "пытались основать теорию ошибок" на идее точности наблюдений, ошибались, потому что точность не является "первоначальным понятием" (1925, с. 74). Соответственно, Леви (1924, с. 77; 1925, с. 284 – 285) признал среднюю квадратическую ошибку лишь как неполную характеристику точности и не согласился с Бьенеме (1853), который отрицал практическую значимость закона Коши. Этот закон, пояснил Леви, доказывает, что нельзя уравнивать наблюдения вне зависимости от распределения их ошибок.

По существу Леви атаковал Лапласа и Гаусса (1823b), которых он (1924, с. 77) также упомянул. "Обманчивость" результатов двух соосновнателей теории ошибок, как он утверждал, "выявилась, когда в 1853 г. Коши привлек внимание" к устойчивым законам и особенно к закону Коши. Возможность надежной оценки точности Леви обусловил устойчивыми законами распределения ошибок (§ 8.5).

**8.3. Средняя квадратическая ошибка.** Леви рассматривал истинные ошибки $\xi_i$, а не уклонения от арифметического среднего и, хоть он этого не сказал, его средняя квадратическая ошибка конечно же равнялась $\sqrt{[\xi\xi]/n}$ . Эта статистика, как он (1925, с. 75) признал, повторяя свое прежнее высказывание (1924, с. 52), соответствует "наиболее простой идее" и ее использование является "довольно естественным". Мало того (1924, с. 74), "представляется, что лучшего параметра действительно нельзя выбрать".

И далее (1925, с. 77): "за неимением лучшего" (очевидно: не зная распределения ошибок) средняя квадратическая ошибка дает "определенное представление о порядке величины ошибки". На с. 61, однако, Леви заявил, что при некоторых положительных $p$ иные оценки типа $\sum|\xi_i|^p/n$ “впрочем” также допустимы. Ни о соображениях Гаусса (§ 5.2.2), ни о возражениях Гельмерта (§ 6.1.3) по этому поводу он не упомянул. Кроме того (с. 78), при нормальном распределении средняя квадратическая ошибка не лучше любого “параметра, определенного иначе”, хотя для распределений, близких к нормальному (с. 78 и 282), почему-то важно применять именно ее. Наконец, Леви (1925, с. 77) сформулировал весьма общее заключение, которое, однако, не имело никакого отношения к делу:

> *Теория, построенная на произвольно введенных аксиомах, не может иметь никакого значения …*

Итак, выборочная дисперсия является удобной оценкой точности, однако возможны и иные статистики и во всяком случае (§ 8.5) без знания закона распределения всесторонняя оценка точности неосуществима, а для некоторых распределений дисперсия просто не существует.

**8.4. Новое понятие точности.** Леви (1924, с. 73) предложил оценивать точность случайной ошибки $\xi$ (лучше сказать: наблюдения, искаженного этой величиной) “параметром”, который указывал бы “порядок величины абсолютной ошибки, которую следует ожидать”. Он (с. 75) заметил, что такой параметр определяется лишь с точностью до произвольного множителя и обосновал свою мысль таким образом (с. 78)

> *Рассматривать понятие параметра точности как интуитивное означает признание, что можно единым числом* [без знания соответствующей плотности, ср. § 8.5] *определять преимущества некоторого метода измерений.*

Почему только метода измерений, а не данных результатов?

При нормальном законе (там же) параметр для выборочного среднего в $\sqrt{n}$ меньше, чем для отдельного наблюдения (1925, с. 280), а его “модуль точности [его вес] $h = 1/\alpha^2$ в $n$ раз больше”. На с. 81 он повторил это выражение, $1/\alpha^2$, но не пояснил его. А как быть, если распределение иное?

**8.5. Устойчивые законы.** Леви определил устойчивые законы в терминах характеристических функций; мы, однако, интересуемся лишь следствием устойчивости (Леви 1924, с. 69; 1925, с. 258): пусть даны независимые и одинаково распределенные ошибки $\xi_i$ и положительные числа $a_i$, $i = 1, 2, \ldots, n$. Если существует такое число $A > 0$, что

$$A^\alpha = a_1^\alpha + a_2^\alpha + \ldots + a_n^\alpha$$

и $[a\xi]/A$ имеет то же распределение, то это распределение устойчиво.

Два условия должны быть дополнительно соблюдены, опять же в качестве следствий (Леви 1924, с. 70; 1925, с. 255). Во-первых, $0 < \alpha \leq 2$; во-вторых, дисперсия устойчивого закона конечна, если и только если $\alpha = 2$.

При $a_i = 1/n$, $A = n^{-(\alpha-1)/\alpha}$ и $[a\xi]/A = \bar{\xi}/A$ имеет то же распределение как каждое $\xi_i$. Пусть $\alpha = 2$, тогда $\bar{\xi}$ распределено так же, как $\xi_i/n$, но для $\alpha = 1$ оказывается, что $\bar{\xi}$ распределено как $\xi_i$. Эти два случая соответствуют нормальному закону и распределению Коши.

Устойчивость важна, потому что (Леви 1925, с. 78; см. также с. 282)

> *Средние … вычисленные при различных системах коэффициентов, не приводят к ошибкам одного и того же типа и, следовательно, их непросто сравнивать с точки зрения точности* [которая, видимо, оценивается "параметрами"] *если этот тип не устойчив.*

Пусть распределение действительно устойчиво. Что тогда? Если $\alpha = 2$, то МНКв[1] сохраняет свое значение (Леви 1925, с. 79): "закон Гаусса является действительно единственным, при котором этот метод применим". Если же $1 < \alpha < 2$, то вес наблюдения $i$ должен быть пропорционален $a_i^{-(\alpha-1)/\alpha}$, где $a_i$ – соответствующий параметр точности (1924, с. 75 – 76; 1925, с. 283). В этом случае (1924, с. 77; 1925, с. 285) наблюдения должны уравниваться по МНКв "с некоторыми видоизменениями". Пусть $a_i = \text{Const}$, так что $\bar{\xi}$ имеет то же распределение, что и $\xi_i n^{-(\alpha-1)/\alpha}$ (т. е., например, что и $\xi_i/n^{1/3}$, если $\alpha = 3/2$). Но что из этого следует? Леви (1924, с. 77) заявил, что надо ввести апостериорные веса, убывающие к краям распределения.

Действительно, раз дисперсия бесконечна, крупные ошибки весьма опасны и их влияние должно быть уменьшено. Но, во-первых, апостериорные веса лишь обеспечивают поправку за асимметрию распределения; во-вторых, Леви никак не развил своей мысли.

Остаются еще два случая, $\alpha = 1$ и $0 < \alpha < 1$. Если $\alpha = 1$, следует выбрать произвольные веса и вычислить обобщенное среднее арифметическое, точность которого, однако, не будет выше, чем у отдельного наблюдения. Но зачем тогда нужно среднее? Не проще ли выбрать любое измерение и отбросить все остальные?

Если же $0 < \alpha < 1$, среднее хуже, чем отдельное наблюдение. Поэтому (Леви 1924, с. 76; 1925, с. 79 и 284)

> *Можно также отбросить в определенной пропорции наибольшие и наименьшие результаты и принять среднее из оставшихся.*

Наиболее просто, продолжал он, оставить лишь половину или треть.

Итак, по Леви распределение ошибок должно быть устойчивым, в противном же случае уравнивание опасно. Но как проверить это условие? И более того: как узнать, будет ли у этого устойчивого закона $\alpha < 1$ или $\alpha > 1$? Впрочем, рекомендация Леви отбросить часть наблюдений (или по крайней мере усомниться в них) может оказаться разумной, см. Элашов и Элашов (1978, с. 233):

*Если у распределения возможны длинные хвосты, преимущества отбрасывания некоторых крайних наблюдений превышают потери*[2].

Эти авторы приводят и соответствующие формальные рекомендации.

Видно, что попытка Леви ввести устойчивые законы в теорию ошибок оказалась мертворожденной. Он, однако, сформулировал некоторые разумные рекомендации и, главное, неожиданно для самого себя существенно продвинул теорию вероятностей. Если же оценивать творчество Леви в целом, то его можно считать сооснователем современной теории вероятностей.

**Примечания**

**1.** Леви очевидно имел в виду среднее арифметическое, а не МНКв вообще. Позднее он (1929, с. 30) отнесся к МНКв (снова: к среднему арифметическому) более благосклонно, утверждая, что его можно применять, если, начиная с какого-то $n$ он “приводит к тем более точным значениям, чем значительнее число $n$”. Леви видимо имел в виду устойчивые законы с параметром $\alpha > 1$, но ведь выполнение этого условия (даже двух условий) трудно или невозможно проверить. И вообще (Золотарев 1984, с. 30 – 31) сравнительно простые выражения для плотностей устойчивых законов известны лишь для $\alpha = 2$, 1 и 1/2.

**2.** Позднее сам Леви (1929. с. 29) высказался в этом же смысле, притом без упоминания устойчивости.

**Основная литература**

Леви (1924; 1925), Шейнин (1995b)

**9. Детерминированная теория ошибок**

**9.1. Планирование эксперимента и предварительное исследование данных.** Мы возвращаемся к соотношению между детерминированной теорией ошибок, планированием эксперимента и предварительным исследованием данных (§§ 0.1 – 0.2).

**9.1.1.** Пусть искомая константа $w$ является функцией $f$ нескольких (например, двух) констант, $u$ и $v$, определенных с погрешностями $du$ и $dv$. Тогда погрешность $dw$ будет равна

$$dw = (\partial f / \partial u)\, du + (\partial f / \partial v)\, dv$$

и, конечно же, следует выбрать более или менее подходящую функцию $f$ (т. е. наилучшие методы и обстоятельства наблюдения).

Так (§ 0.1), при засечке пункта $C$ со станций $A$ и $B$ требуется определить оптимальную форму треугольника $ABC$. В этом примере $dA$ и $dB$ – ошибки измерения углов, которые повлияют на координаты $x$ и $y$ искомого пункта:

$$dx = (\partial f_1 / \partial A)\, dA + (\partial f_1 / \partial B)\, dB,\ dy = (\partial f_2 / \partial A)\, dA + (\partial f_2 / \partial B)\, dB.$$

Если несколько расширить понятие планирования эксперимента, то подобные исследования можно будет включить в него.

**9.1.2.** При обработке наблюдений необходимо уменьшить влияние систематических ошибок на их результаты. Этого можно частично добиться вероятностными методами (§§ 7.2 – 7.3), а частично – сравнительно простыми приемами, почти независимо от теории вероятностей, т. е. методами предварительного исследования данных, ср. § 9.4.4.

Вряд ли имеет смысл систематизировать эти приемы, но эту сторону детерминированной теории ошибок полезно иметь в виду.

**9.2. Восемнадцатый век.** Древние астрономы хорошо знали, что наблюдения искажаются ошибками и что их следует выполнять при наилучших условиях (§§ 1.3 и 1.6.3), и они также имели некоторое представление о различии между систематическими и случайными ошибками (§ 1.4). То же тем более верно по отношению к ученым XVI – XVII веков, см. также Прим. 13 гл. 2.

Впрочем, по-настоящему детерминированная теория ошибок возникла в XVIII в.; о Лапласе мы, однако, будем говорить в § 9.3.

Котс (посмертная публ. 1722) решил 28 простых задач, о связи между дифференциалами различных элементов прямолинейных и сферических треугольников. Тем самым он дал возможность устанавливать влияние погрешностей измерений на косвенно определяемые элементы. Его сочинение стало широко известно, и, например, Кондамин (1751, с. 91) заявил, что применял "теорию месье Котса".

По мнению Ламберта (1765a, § 321), который не сослался ни на кого, изучение функций наблюденных величин составляло предмет *теории последствий* (Theorie der Folgen) [ошибок]. В §§ 340 – 426 он установил выгоднейшие формы стандартных геодезических фигур, хотя и не цепей триангуляции.

Даниил Бернулли (§ 2.8) определил систематические и случайные ошибки, хотя по необходимости лишь в узком смысле. Тем самым он всё-таки сделал важный и необходимый шаг в изучении ошибок наблюдения.

Майер и Бошкович исследовали астрономические приборы и вывели дифференциальные уравнения, соединяющие их погрешности с соответствующими неизбежно возникающими ошибками измерений (например, в моменте прохождения звезды через меридиан). Провербио (1988), который описал работы обоих ученых, заметил, что формулы, подобные выведенным ими, позволяли полностью положиться на предварительную юстировку прибора; мы бы сказали, – позволяли вырабатывать наилучшие программы наблюдений.

Позже 1752 г. Майер (который умер в 1762 г.) кроме того изобрел повторительный теодолит, вошедший в употребление к концу того же века (Форбс 1974, с. 235; Кларк 1880, с. 13). Работая с этим прибором, наблюдатель мог несколько раз ($n$ раз) откладывать измеряемый угол на лимбе, сопровождая это лишь двумя отсчетами (в самом начале измерения, по левому, например, сигналу, и в самом конце, по правому). После этого величину кратного угла, разумеется, делили на $n$.

Ошибка отсчета была намного крупнее ошибки визирования, но ее влияние уменьшалось в $n$ раз и этот параметр $n$, хоть и не был вполне произволен, позволял уравнивать действие обеих указанных ошибок и тем самым уменьшать общую ошибку и добиваться лучшего выполнения предпосылок ЦПТ; см. соответствующее мнение Лапласа в § 3.6.1.

Детерминированный подход начал применяться и при уравнивании наблюдений. Вспомним, что Майер (§ 2.2) придумал прием для лучшего определения своего основного неизвестного и что Гаусс (правда, только в своей переписке и лишь при рядовом исследовании) поступил аналогично, не став применять МНКв.

**9.3. Лаплас**

**9.3.1.** Многие сочинения Лапласа содержат утверждения о наилучших программах наблюдения, о влиянии ошибок на конечные результаты и т. д., а в некоторых случаях и о планировании эксперимента в целом. Так, он (1821) описал с этой точки зрения метод определения орбит комет, а его ранняя работа (1784) была посвящена общим соображениям на ту же тему.

**9.3.2.** Более интересны *Дополнения 2* (1818) и *3* (прим. 1819) к *Аналитической теории вероятностей*. В них, исходя из ошибок угловых измерений, Лаплас исследовал точность длины дуги меридиана, вычисленной при помощи цепи триангуляции. Во второй работе он дополнительно рассмотрел цепи, состоящие из конгруентных равнобедренных треугольников, а затем оценил точность тригонометрического нивелирования (основанного на измерениях зенитных расстояний).

Система обозначений, принятая Лапласом, неудачна, и пояснения в *Дополнении 2* недостаточны, да и вообще читатель, знакомый с уравнительными вычислениями по Гауссу, т. е. не исходящими из предположения о нормальном распределении, будет явно разочарован выкладками автора. И тем не менее именно Лаплас первым исследовал точность цепей триангуляции и таким образом сделал дальнейший важный шаг в изучении геодезических построений.

**9.3.3.** В § 1.9 мы упомянули Кеплера в связи с методом минимакса, а сейчас мы коснемся соответствующей работы Лапласа (1792, с. 506), который изучал определение фигуры Земли:

> *Установленный эллипс служит для выяснения, находится ли эллиптическая фигура в границах ошибок наблюдения; но это не тот эллипс, на который с наибольшим правдоподобием указывают* [измеренные] *градусы* [меридианов].

Это утверждение вполне соответствует нашему предположению о подходе Кеплера. В своей *Аналитической теории* Лаплас (1812, с. 351) заявил, что "после многих примеров" он определил, что результаты методов минимакса и наименьших квадратов лишь незначительно отличаются друг от друга. Мы, впрочем, не беремся судить, так ли это действительно. На алгоритме вычислений по методу минимакса (там же, см. § 24, к которому относится и упомянутая выше с. 351) мы не останавливаемся.

Позднее метод минимакса применялся для приближения нелинейных (в том числе не алгебраических) функций. В частности, Понселе и Чебышев решили при его помощи важную задачу из теории механизмов, а именно преобразование кругового движения в прямолинейное, а намного раньше Эйлер применил его же для целей картографии (Гусак 1961).

**9.4. Гаусс и Бессель.** Гиппократ, Браге и Брадлей, как справедливо считается, были наблюдателями высшего класса, но Гаусс и Бессель оказались зачинателями стадии экспериментальной науки, которую можно, пожалуй, назвать классической. Не без основания Ньюком (Шмайдлер 1984, с. 32 – 33) упомянул "немецкую школу практической астрономии", хоть и назвал только Бесселя. Основная идея этой школы, как он заявил, состояла в том, что каждый прибор обвинялся

*Во всех возможных недостатках и не оправдывался до тех пор, пока не доказывал себя безупречным во всём. Бессель усовершенствовал приемы определения возможных погрешностей приборов с изобретательностью и точностью геометрического метода …*

Но и Гаусс, и Бессель исследовали и методы наблюдения.

**9.4.1.** Гаусс изучил основные систематические ошибки, возникающие в процессе наблюдений, а именно вызванные боковой рефракцией, неточным делением лимбов теодолита и методом измерения углов повторительным теодолитом, (*методом повторения*) см. Гаусс (1958), где в переводе из его Трудов, тт. 4 и 9, можно найти его соответствующие мнения и рекомендации из одной рецензии 1830 г. и переписки. .

Напомним также (§ 5.4.1), что Гаусс не придерживался единой программы наблюдений и добавим, что он изобрел и с успехом применял гелиотроп (солнечное зеркало), см. там же.

**9.4.2.** Особого внимания заслуживает метод Гаусса определения разности между двумя примерно равными весами двух тел ($A$ и $B$), поскольку он существенно усовершенствовал метод Борда (Гельмерт 1872, с. 47 – 49). Мало того, Пукельсхейм (1993, с. 427) указал на близость метода Гаусса современным понятиям взвешивания, сам же Гаусс описал его только в письмах Шумахеру 1836 и 1839 гг. (1861/1975, № 3, с. 99 – 101, 268, 272 – 275 и 330 – 333).

Борда взвешивал $A$ и $B$ по-отдельности и уравнивал их веса добавлением к ним весов $a_i$ или $b_i$, добиваясь одного и того же отсчета $T$ на весах и учитывая при этом постоянную ошибку измерения $c$ и другую возможную ошибку $w$, пропорциональную времени. Вот его уравнения:

$$T = A + a_1 + c + \varepsilon_1, \qquad T = B + b_2 + c + 2w + \varepsilon_3,$$
$$T = B + b_1 + c + w + \varepsilon_2, \quad T = A + a_2 + c + 3w + \varepsilon_4,$$

в которых $\varepsilon_i$ – случайные ошибки взвешивания. Исключая $c$ и $w$, Борда получил

$$A - B = (1/2)[(b_1 + b_2) - (a_1 + a_2)] + (1/2)[(\varepsilon_2 + \varepsilon_3) - (\varepsilon_1 + \varepsilon_4)].$$

Гаусс, однако,взвешивал $A$ и $B$ одновременно:

$$B = A + a_1 + c + \varepsilon_1, \qquad A = B + b_2 + c + 2w + \varepsilon_3,$$
$$A = B + b_1 + c + w + \varepsilon_2, \quad B = A + a_2 + c + 3w + \varepsilon_4.$$

Таким образом,

$$A - B = (1/4)\ [(b_1 + b_2) - (a_1 + a_2)] + (1/4)[(\varepsilon_2 + \varepsilon_3) - (\varepsilon_1 + \varepsilon_4)]$$

и точность результата оказывалась вдвое выше, поскольку исключалось дополнительное неизвестное $T$.

**9.4.3.** Бессель вместе с Байером выполнили триангуляцию Восточной Пруссии и он, уже самостоятельно, описал эту работу (1838b). Там он обсудил измерение базисов при помощи мерных жезлов, исследование теодолитов и методы геодезических и астрономических наблюдений.

Первая из перечисленных тем была, видимо, менее всего изучена и следует упомянуть специальное исследование Бесселя (1839). Мерный жезл длиной в несколько футов опирается на подставки в двух своих точках, расположенных, естественно, симметрично относительно его середины.

Жезл изгибается под действием своего собственного веса, и Бессель при помощи составленных и решенных им дифференциальных уравнений определил, где именно должны быть расположены опорные точки, чтобы жезл укоротился меньше всего. Для современного инженера-строителя это исследование быть может и не интересно, но в свое время оно безусловно было новинкой.

**9.4.4.** Бессель (1823) установил существование личного уравнения (личной ошибки) астронома. Оказалось, что момент прохождения звезды через крест нитей окуляра астрономического прибора регистрируется каждым наблюдателем по-своему, причем в течение сравнительно короткого времени разность таких моментов, зарегистрированная двумя астрономами, оставалась примерно постоянной.

Бессель сослался на Маскелайна, который в 1796 г. уволил своего помощника, поскольку обнаружил, что их наблюдения систематически разнятся на $0.^{s}8$. Далее он исследовал разности зарегистрированных моментов для нескольких пар астрономов и пришел к указанному выше заключению. Так как наблюдатели не могли работать одновременно, приходилось учитывать поправку за ход хронометра (который Бессель определял отдельно).

Однако, в одном случае, в котором этот ход определялся из самих наблюдений, выводы Бесселя просто никуда не годились. Вот полученные им уравнения; наблюдатели – сам Бессель и Вальбек, числа указаны в секундах времени, $x$ – разность “Бессель” минус “Вальбек”, $y_i$ – поправки за ход хронометра:

$$x + y_1 = 1.93,\ -x + y_1 = -0.36;\ -x + y_2 = -0.97,\ x + y_2 = 1.00;$$
$$x + y_3 = 1.10,\ -x + y_3 = -0.92;\ -x + y_4 = -0.09,\ x + y_4 = 1.96.$$

Решив каждую пару уравнений по-отдельности, Бессель получил

$x = 1.145$; $0.985$; $1.010$; и $1.025$

и указал (с. 220), что среднее арифметическое, 1.041, “вряд ли может быть ошибочным на величину в несколько сотых долей секунды”.

Этот вывод неприемлем, и, мало того, значения $y_i$, которые Бессель вообще не указал, значительно отличались друг от друга:

$y_i = 0.785$; $0.015$; $0.090$; и $0.935$.

Даже этого мало: эти величины не были независимыми; так, $y_1$ относилось к периоду 16-го – 17-го декабря 1820 г., а $y_2$ – к 17-му и 19-му декабря.

Приходится добавить, что Бесселю были присущи неверные суждения (см. § 6.1.1). Кроме того, мы (2000b) обнаружили 33 ошибки в проделанных им арифметических и простейших алгебраических действиях. Не будучи существенными, они подрывают веру в надежность его более сложных вычислений.

**9.5. Гельмерт**

**9.5.1.** Фактически следуя за Лапласом (§ 9.3.2), Гельмерт (1868) исследовал точность различных геодезических построений и изучил программы наблюдений с точки зрения не существовавшей в то время теории линейного программирования. Именно, он (с.1 и 60 в отдельной публикации) поставил задачу либо лостичь наибольшего веса результатов при заданном объеме работы, либо заданных весов при наименьшем ее объеме. В частности, он (с. 59) заметил, что некоторые углы ромбической базисной сети[1] вообще не следует измерять. Этот правильный вывод был в некоторой степени осуществлен (быть может независимо от него); полностью отказываться от измерения каких-либо углов нельзя ввиду необходимости контроля полевых наблюдений.

Уравнения, возникающие в геодезических сетях, либо линейны, либо линеазируются перед уравниванием, и в принципе Гельмерт мог бы создать элементы теории линейного программирования, но этого не произошло. Его исследование продолжили Шрейбер (1882) и Брунс (1882; 1886). Современным трудом является работа Графаренда и Харланда (1973), а промежуточной – статья Фридриха (1937). Напомним (§§ 2.5.1 и 5.1.1), что метод Бошковича решения систем линейных уравнений равнозначен выбору базисного решения в соответствующей задаче теории линейного программирования и что Гаусс и Лаплас знали требуемую здесь важную теорему этой теории.

**9.5.2.** Гельмерт (1886, с. 1 и 68) был также первым, кто предложил упростить уравнивание большой сети триангуляции временной заменой ее частей геодезическими линиями. Сеть, которую он уравнивал, была сложной и в бóльшей своей части устарела, и Гельмерт в первую очередь хотел придумать разумный приближенный метод ее обработки.

Много позже Красовский использовал идею Гельмерта для уравнивания стройной системы советской триангуляции, которая состояла из звеньев с базисами и двусторонними азимутами Лапласа на концах каждого из них. Эти звенья в предварительном порядке уравнивались по-отдельности (базисы и азимуты исправлять не надо было, поскольку они были намного точнее угловых измерений) и заменялись геодезическими линиями; они-то и уравнивались совместно, а затем от них вновь переходили к звеньям треугольников и окончательно уравнивали их, снова по-отдельности.

Вот что было не совсем справедливо сказано о Гельмерте (Закатов 1950, с. 369 и 371):

> *Применить обычный метод уравнивания путем совместного решения всех возникающих в сети условных уравнений … не представляется возможным вследствие чрезвычайной громоздкости работы. Кроме того, как указывает Ф. Н. Красовский, … имеются сомнения в том, что результаты такого уравнивания будут наилучшими. … Указанный путь уравнивания …был предложен* [им. Он видоизменил] *метод Гельмерта, который в силу чрезвычайной сложности и громоздкости не мог быть применен для уравнивания такой обширной … сети …*

А. А. Изотов, ближайший помощник Красовского, на лекции примерно 1950 г., на которой мы присутствовали, пояснил, что при уравнивании без применения геодезических (будь это возможно), систематические ошибки *свободно гуляли бы по всей сети*.

**Примечания**

**1.** "Ромбом" служил четырехугольник с весьма неравными диагоналями. Ввиду очевидных трудностей, непосредственно измерялась меньшая из них, бóльшая же после ее вычисления служила стороной цепи триангуляции. Базисные сети ввел Шверд (1822).

## Библиография

AHES = *Arch. Hist. Ex. Sci.*
AN = *Astron. Nachr.*
OC = *Oeuvr. Compl.*
ZfV = *Z. f. Vermessungswesen*

**Архимед**, см. также **Archimedes** (1932), *Исчисление песчинок* (*Псаммит*). М. – Л.

**Берви Н. В.** (1899), Определение вероятнейшего значения измеряемого объекта помимо постулата Гаусса. *Имп. Моск. Общество любителей естествозн., антропологии и этнографии*, отд. физич. наук, т. 10, вып. 1, с. 41 – 45.

**Бернулли Д.**, см. также **Bernoulli D.** (1778, латин.), Наиболее вероятный выбор из нескольких не согласующихся друг с другом наблюдений и т. д. В книге Шейнин (2006а, с. 237 – 254, 263 – 267).

**Бессель Ф. В.**, см. также **Bessel F. W.** (1823, нем.), Личное уравнение при наблюдениях прохождения звезд. В книге автора (1961, с. 219 – 225).

--- (1834, нем.), Некоторые соображения по поводу наблюдений способом повторений. Там же, с. 200 – 218.

--- (1838a, нем.), Исследование о вероятности ошибок наблюдений. Там же, с. 226 – 258.
--- (1838b, нем.), Градусное измерение в Восточной Пруссии и т. д. Там же, с. 99 – 186 (частичный перевод).
--- (1839, нем.), Влияние силы тяжести на фигуру жезла и т. д. Там же, с. 187 – 199.
--- (1961), *Избранные геодезические сочинения. Высшая геодезия и способ наименьших квадратов.* Редактор Г. В. Багратуни. М.
**Бируни А. Р**., см. также **Al-Biruni R.** (1966), *Определение границ мест для уточнения расстояний между населенными пунктами. Избр. произв*., т. 3. Ташкент.
--- (1983), Об отношениях между металлами и драгоценными камнями по объему. В книге *Из истории физ.-мат. наук на средневековом Востоке. Научное наследство*, т. 6. М., с. 141 – 160. См. там же, с. 15 – 140, трактат Ал-Хазини, третья книга которого (с. 52 – 75) является, по ред. прим., переработкой Бируни.
**Бомфорд Г.**, см. также **Bomford G.** (1952, англ.), *Геодезия*. М. Перевод: О. Б. Шейнин.
**Боярский А. Я.,** редактор (1970), *Введение в теорию порядковых статистик*. М.
**Буняковский В. Я.** (1846), *Основания математической теории вероятностей.* СПб.
**Васильев А. В.** (1885), *Теория вероятностей*. Казань. Литография.
**Галилей Г.**, см. также **Galilei G.** (1632, итал.), *Диалог о двух главнейших системах мира.* М. – Л., 1948. Перевод: А. Н. Долгов.
**Гаусс К. Ф.**, см. также **Gauss C. F.** (1809a, нем.), Теория движения, авторское сообщение. В книге автора (1957, с. 150).
--- (1809b, латин.), Теория движения и т. д. Там же, с. 89 – 109.
--- (1811, латин.), Исследование об эллиптических элементах Паллады и т. д. Там же, с. 111 – 120.
--- (1816, нем.), Определение точности наблюдений. Там же, с. 121 – 128.
--- (1821, нем.), Теория комбинаций наблюдений и т. д., часть 1-я, авторское сообщение. Там же, с. 141 – 144.
--- (1822, нем.), Приложение теории вероятностей к одной задаче практической геометрии. Там же, с. 129 – 133.
--- (1823a, нем.), Теория комбинаций наблюдений и т. д., часть 2-я, авторское сообщение. Там же, с. 144 – 147.
--- (1823b, латин.), Теория комбинаций наблюдений и т. д., части 1 и 2. Там же, с. 17 – 57.
--- (1826, нем.), Дополнение к теории комбинаций наблюдений и т. д., авторское сообщение. Там же, с. 147 – 150.
--- (1828, латин.), Дополнение к Теории комбинаций наблюдений и т. д. Там же, с. 59 – 88.
--- (1957), *Избранные геодезические сочинения*, т. 1. *Способ наименьших квадратов*. Редактор Г. В. Багратуни. М.
--- (1958), *Избранные геодезические сочинения,* т. 2. *Высшая геодезия*. Редактор Г. В. Багратуни. М.
**Гегель Г. В. Ф.** (1812, нем.), *Наука логики*. *Соч*., т. 5. М., 1937.

**Гельмерт Ф. Р.**, см. также **Helmert F. R.** (1872, нем.), *Уравновешивание по способу наименьших квадратов* и т. д. М., 1914. Неполный перевод (Л. А. Сопоцько) с издания 1907 г.

**Гильберт Д.** (1901, нем.), *Проблемы Гильберта*. М., 1969.

**Гнеденко Б. В., Шейнин О. Б.** (1978), Теория вероятностей. Глава в книге *Математика XIX века*, т. 1. Редакторы А. Н. Колмогоров, А. П. Юшкевич. М., с. 184 – 240.

**Гусак А. А.** (1961), Предыстория и начало развития теории приближения функций. ИМИ, т. 14, с. 289 – 348.

**Декарт Р.**, см. также **Descartes R.** (1637, франц.), *Рассуждение о методе. Избр. Произв.* М., 1953.

**Диксон У.** (1962, англ.), Отбраковка сомнительных наблюдений. В книге Боярский (1970, с. 274 – 307).

**Закатов П. С.** (1950), *Курс высшей геодезии*. М., 1953, 1964.

**Золотарев В. М.** (1984), *Устойчивые законы и их применения.* М.

**Идельсон Н. И.** (1947), *Способ наименьших квадратов и теория математической обработки наблюдений.* М.

**Кемниц Ю. В.** (1957), О функции распределения ошибок измерений. *Геод. и картография*, № 10, с. 21 – 29.

**Кетле А.**, см. также **Quetelet A.** (1836, франц.), *Человек и развитие его способностей или опыт общественной физики.* СПб, 1865.

**Колмогоров А. Н.** (1931), Метод медианы в теории ошибок. *Математич. Сб.*, т. 38, № 3 – 4, с. 47 – 49.

--- (1946), К обоснованию метода наименьших квадратов. *Успехи математич. наук*, т. 1, № 1, с. 57 – 71.

**Колмогоров А. Н. и др.** (1947), Одна формула из теории метода наименьших квадратов. *Изв. АН СССР*, сер. математич., т. 11, с. 561 – 566.

**Корфельд М.** (1955), К теории ошибок. *Докл. АН СССР*, т. 103, № 2, с. 213 – 214.

**Крамер Г.** (1946, англ.), *Математические методы статистики*. М., 1948.

**Курно О.** (1843, франц.), *Основы теории шансов и вероятностей*. М., 1970.

**Лаплас П. С.**, см. также **Laplace P. S.** (1796, франц.), *Изложение системы мира.* СПб, 1861.

--- (1812, франц.), О вероятности ошибок среднего результата большого числа наблюдений и о наиболее благоприятных средних результатах (глава 4 из *Аналитической теории вероятностей* автора). В книге Шейнин (2007, с. 94 – 131).

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

**Лежандр А. М.**, см. также **Legendre A. M.** (1814, франц.), Метод наименьших квадратов для определения вероятнейшего среднего из результатов различающихся наблюдений. Перепечатка мемуара 1805 г. с измененным заглавием. В книге Шейнин (2007, с. 73 – 76).

**Леман Э.** (1959, англ.), *Проверка статистических гипотез.* М., 1979.

**Линник Ю. В.** (1952), Замечания по поводу классического вывода закона Максвелла. *Докл. АН СССР*, т. 85, № 6, с. 1251 – 1254.

--- (1958), *Метод наименьших квадратов* и т. д. М., 1962.
**Линник Ю. В. и др.** (1951), Очерк работ А. А. Маркова по теории чисел и теории вероятностей. В книге Марков (1951, с. 614 – 640).
**Лурье С. Я.** (1934), Приближенные вычисления в древней Греции. *Архив истории науки и техн.*, вып. 4, с. 21 – 46.
**Майстров Л. Е.** (1964), Элементы теории вероятностей у Галилея. *Вопр. истории естеств. и техники*, вып. 16, с. 94 – 98.
--- (1967), *Теория вероятностей. Исторический очерк*. М.
**Мальцев А. И.** (1947), Замечание к работе Колмогоров и др. (1947). *Изв. АН СССР*, сер. математич., т. 11, с. 567 – 578.
**Марков А. А.** (1899), Закон больших чисел и способ наименьших квадртаов. В книге автора (1951, с. 231 – 251).
--- (1900), *Исчисление вероятностей*. Посл. изд.: 1908, 1913 и посм. 1924.
--- (1951), *Избранные труды.* Без места.
**Менделеев Д. И.** (1860), О сцеплении некоторых жидкостей. *Соч.*, т. 5. М. – Л., 1947, с. 40 – 55.
--- (1872), О сжимаемости газов. *Соч.*, т. 6. М. – Л., 1939, с. 128 – 171.
--- (1875), Ход работ по возобновлению прототипов. *Соч.*, т. 22. Л. – М., 1950, с. 175 – 213.
--- (1887), Исследование водных растворов по удельному весу. *Соч.*, т. 3. М. – Л., 1934, с. 3 – 468.
--- (1895), О весе определенного объема воды. *Соч.*, т. 22, с. 105 – 171.
**Милл Дж. С.** (1843, англ.), *Система логики*. СПб, 1914. Перевод с изд. 1879 г.
**Ондар Х. О.**, редактор (1977), *О теории вероятностей и математической статистике. Переписка А. А. Маркова и А. А. Чупрова*. М.
**Петров В. В.** (1954), О методе наименьших квадратов и его экстремальных свойствах. *Успехи математич. наук*, т. 9, с. 41 – 62.
**Прения** (1903), *Прения между академиками в заседаниях 1-го отделения Академии наук* [в связи с представлением Б. Б. Голицина к избранию ординарным академиком]. СПб.
**Птолемей К.**, см. также **Ptolemy** (1998), *Альмагест*. М. Перевод: И. Н. Веселовский.
**Пуанкаре А.**, см. также **Poincaré H.** (1896, франц.), *Теория вероятностей*. Ижевск, 1999.
**Розенталь И. С., Соколов В. С.** (1956), *Учебник латинского языка.* М.
**Сархан С., Гринберг Б.** (1962, англ.), Оптимальные линейные несмещенные оценки для некоторых симметричных распределений. В книге Боярский (1970, с. 351 – 359).
**Субботин М. Ф.** (1956), Астрономические и геодезические работы Гаусса. В мемориальном сборнике *К. Ф. Гаусс*. М., с. 243 – 310.
**Уиттекер Э., Робинсон Г.**, см. также **Whittaker E. T., Robinson G.** (1924, англ.), *Математическая обработка результатов наблюдений.* М., 1935.

**Усов Н. А.** (1867), Замечание по поводу теоремы П. Л. Чебышева. *Математич. Сб.*, т. 2, с. 93 – 95 второй пагинации.

**Фишер Р. А.**, см. также **Fisher R. A.** (1925, англ.), *Статистические методы для исследователей*. М., 1958.

**Цингер** [**В. Я.**] (1862), *Способ наименьших квадратов.* М. Диссертация.

**Чеботарев А. С.** (1955), *Геодезия*, ч. 1. М.

--- (1958), *Способ наименьших квадратов* и т. д. М.

**Чебышев П. Л.** (лекции 1879/1880), *Теория вероятностей*. М. – Л., 1936.

**Шевченко М. Ю.** (1988), Звездный каталог Птолемея. *Историко-астрономич. исследования*, вып. 20, с. 167 – 186.

**Шейнин О. Б.**, см. также **Sheynin O. B.** (1965), О работах Р. Эдрейна в теории ошибок. ИМИ, т. 16, с. 325 – 336.

--- (1967), К истории уравнивания косвенных наблюдений. *Изв. вузов. Геод. и аэрофотосъемка*, № 3, с. 25 – 32.

--- (1990), Отзыв А. А. Маркова об одной статье Б. Б. Голицина. ИМИ, т. 32 – 33, с. 451 – 455.

--- (1991, англ.), Понятие случайности от Аристотеля до Пуанкаре. ИМИ, т. 1 (36), № 1, 1995, с. 85 – 105.

--- (2000а), История теории ошибок. ИМИ, т. 5 (40), с. 310 – 332.

--- (2005а), *Теория вероятностей. Исторический очерк*. Берлин.

--- (2006а), *Хрестоматия по истории теории вероятностей и статистики.* Берлин.

--- (2007), *Вторая хрестоматия по истории теории вероятностей и статистики*. Берлин.

**Шейнин О. Б., Майстров Л. Е.** (1972), Теория вероятностей. Глава в книге *История математики с древнейших времен до начала XIX столетия*, т. 3. Редактор А. П. Юшкович. М., с. 126 – 152.

**Эйлер Л.,** см. также **Euler L.** (1778, латин.; 1961, англ.), Замечания к предшествующему рассуждению прославленного Д. Бернулли. В книге Шейнин (2006а, с. 254 – 267).

**Этьен**, см. также **Estienne J. E.** (1890, франц.), Исследование ошибок наблюдения. *Артилл. Ж.*, 1895, № 8, с. 703 – 723. Перевод: С. Д [ельвиг].

**Юл Дж. У., Кендалл М. Дж.** (1937, англ.), *Теория статистики*. М., 1960.

**Ярошенко С. П.**, см. также **Yarochenko S.** (1893а), К теории способа наименьших квадратов. *Зап. Новоросс. Унив*., т. 58, с. 193 – 208 второй пагинации.

**Aaboe A., De Solla Price D. S.** (1964), Qualitative measurements in antiquity. В книге *L'aventure de la science. Mélanges A. Koyré*, t. 1. Paris, pp. 1 – 20.

**Abbe C.** (1871), Historical note on the method of least squares. *Amer. J. Sci. Arts*, vol. 1, pp. 411 – 415. Перепечатка в книге Stigler (1980, vol. 1).

**Abbe E.** (1863), Über die Gesetzmässigkeit in der Verteilung der Fehler bei Beobachtungsreihen. *Ges. Abh.*, Bd. 2. Hildesheim, 1989, pp. 55 – 81.

**Adrain R.** (1808), Research concerning the probabilities of the errors which happen in making observations. Перепечатка в книге Stigler (1980, vol. 1).
--- (1818a), Investigation of the figure of the Earth and of the gravity in different latitudes. Там же.
--- (1818b), Research concerning the mean diameter of the Earth. Там же.
**Airy G. B.** (1861), *On the Algebraical and Numerical Theory of Errors of Observations* etc. Cambridge.
**Al-Biruni R.** (1967), *The Determination of the Coordinates of Position* etc. Beirut.
**Archimedes** (1925), Die Sandzahl. В книге автора *Über schwimmende Körper und Sandzahl.* Leipzig, pp. 67 – 82.
**Aristarchus** (1959), On the sizes and distances of the Sun and the Moon. В книге Heath T. *Aristarchus of Samos.* Oxford, pp. 353 – 414.
**Barnett V., Lewis T.** (1978), *Outliers in Statistical Data*. Chichester, 1984.
**Benjamin M.** (1910), Newcomb. В книге *Leading American Men of Science*. New York, pp. 363 – 389.
**Berggren J. L.** (1991), Ptolemy's maps of Earth and the heavens: a new interpretation. AHES, vol. 43, pp. 133 – 144.
**Bernoulli D.** (не позднее 1769, латин.), The most probable choice between several discrepant observations etc. В книге *Festschrift for Lucien Le Cam*. Редактор D. Pollard и др. New York, 1997, pp. 358 – 367.
--- (1770 – 1771), Mensura sortis. *Werke*, Bd. 2. Basel, 1982, pp. 326 – 360.
--- (1778, латин.), The most probable choice … *Biometrika*, 1961, vol. 48, pp. 1 – 18. Перепечатка в книге Pearson E. S., Kendall (1970, pp. 157 – 167).
--- (1780), Specimen philosophicum de compensationibus horologicis. *Werke*, Bd. 2, pp. 376 – 390.
**Bernoulli, J.** (рукопись; частичная публ. 1975), [Aus den Meditationes]. *Werke*, Bd. 3. Basel, pp. 21 – 89.
**Bernoulli Johann III** (1789), Milieu. *Enc. méthodique. Dict. enc. math.*, t. 2. Paris, pp. 404 – 409.
**Bertrand J. L. F.** (1855), Sur la méthode des moindres carrés. *C. r. Acad. Sci. Paris*, t. 40, pp. 1190 – 1192.
--- (1888a), Sur l'évaluation a posteriori de la confiance méritée par la moyenne. Там же, t. 106, c. 887 – 891.
--- (1888b), Sur l'erreur à craindre dans l'évaluation des trois angles d'un triangle. Там же, c. 967 – 970.
--- (1888c), Sur les conséquences de l'égalité acceptée entre la valeur vraie d'un polynôme et sa valeur moyenne. Там же, c. 1259 – 1263.
--- (1888d), *Calcul des probabilités*. Второе изд. 1907. Перепечатки: Нью Йорк, 1970, 1972.
**Bessel F. W.** (1816), Untersuchungen über die Bahn des Olbersschen Kometen. *Abh. Preuss. Akad. Wiss.* [Berlin], math. Kl., 1812 – 1813, pp. 119 – 160.
--- (1818) *Fundamenta astronomiae.* Königsberg.

--- (1820), Beschreibung des auf des Königsberger Sternwarte. *Astron. Jahrb*. за 1823. Berlin, pp. 161 – 168.

--- (1826), Methode die Thermometer zu berichtigen. В книге автора (1876, Bd. 3, pp. 226 – 233).

--- (прочтено 1832), Über den gegenwärtigen Standpunkt der Astronomie. В книге автора *Populäre Vorlesungen*. Hamburg, 1848, pp. 1 – 33.

--- (1833), Письмо Дж. Б. Эйри. В книге автора (1876, Bd. 3, pp. 462 – 465).

--- (1876), *Abhandlungen*, Bde 1 – 3. Leipzig.

**Bienaymé I. J.** (1852), Sur la probabilité des erreurs d'après la méthode des moindres carrés. *J. math. pures appl.*, sér. 1, t. 17, pp. 33 – 78.

--- (1853), Considérations à l'appui de la découverte de Laplace. *C. r. Acad. Sci. Paris*, t. 37, pp. 309 – 324. Перепечатка (1867): *J. math. pures appl.*, sér. 2, t. 12, pp. 158 – 176.

**Biermann K.-R.** (1956), Spezielle Untersuchungen zur Kombinatorik durch G. W. Leibniz. *Forschungen u. Fortschritte*, Bd. 30, pp. 169 – 172.

--- (1966), Über die Beziehungen zwischen Gauss und Bessel. *Mitt. Gauss-Ges. Göttingen*, Bd. 3, pp. 7 – 20.

**Biot J. B.** (1811), *Traité élémentaire d'astronomie physique*, t. 2. Paris – St. Petersbourg. Второе изд.

--- (1829), Sur la figure de la Terre. *Mém. Acad. Sci. Paris*, t. 8, pp. 1 – 56.

**Bomford G.** (1952), *Geodesy*. Oxford, 1962, 1971.

**Bopp K.** (1924), *Euler's und Lambert's Briefwechsel. Abh. Preuss. Akad. Wiss.*, phys.-math. Kl., No. 2. Весь выпуск.

**Boscovich R. J.** (1757), De litteraria expeditione par pontificam ditionem. Перепечатано с сербским переводом в книге Cubranic (1961).

--- (без даты, не опубл.), De calculo probabilitatum. Хранится в Boscovich Archive, Dept Rare Books & Sp. Coll., Univ. California Library, MS No. 62.

--- (1758, латинск.), *Theory of Natural Philosophy.* Cambridge, Mass., 1966. Перевод с изд. 1763 г.

**Boyle R.** (посм. публ. 1772), A physico-chymical essay. *Works*, vol. 1. Sterling, Virginia, 1999, pp. 359 – 376.

**Bradley J.** (1750), A letter … concerning the apparent motion observed in some of the fixed stars. В книге Rigaud (1832/1972, pp. 17 – 41).

**Bru B.** (1988), Laplace et la critique probabiliste des mesures géodésiques. В книге Lacombe H., Costabel P., *La figure de la Terre*. Paris, pp. 223 – 244.

**Bruns H.** (1882), Über eine Minimumsaufgabe. *Math. Annalen*, Bd. 20, pp. 455 – 460.

--- (1886), Über eine Aufgabe der Ausgleichung. *Abh. Sächs. Ges. Wiss.*, math.-phys. Kl., Bd. 13, No. 7, pp. 517 – 563.

**Caspar M., von Dyck W.** (1930), *Kepler in seinen Briefen*, Bd. 2. München – Berlin.

**Cauchy A. L.** (1853a), Sur l'évaluation d'inconnues déterminées par un grand nombre d'équations approximatives. OC, sér. 1, t. 12. Paris, 1900, pp. 36 – 46.

--- (1853b), Sur la nouvelle méthode d'interpolation comparée à la méthode des moindres carrés. Там же, с. 68 – 79.

**Chakrabarti C. G.** (1989), Gauss' principle and statistical thermodynamics. *Appl. Math. Lett.*, vol. 2, pp. 121 – 123.

**Chatterjee S. K.** (2003), *Statistical Thought: a Perspective and History.* Oxford.

**Chauvenet W.** (1863), *Manual of Spherical and Practical Astronomy*, vols 1 – 2. New York, 1960.

**Clarke A. R.** (1880), *Geodesy*. Oxford.

**Condamine C. M. de la** (1751), *Mesure des trois premièrs dégrés du méridien.* Paris.

**Condorcet M. J. A. N. Caritat de** (1805), Éléments du calcul des probabilités. В книге автора *Sur les éléction et autres textes*. Paris, 1986, pp. 483 – 623.

**Coolidge J. L.** (1926), R. Adrain and the beginnings of American mathematics. *Amer. Math. Monthly*, vol. 33, No. 2, pp. 61 – 76.

**Corry L.** (1997), Hilbert and the axiomatization of physics. AHES, vol. 51, pp. 83 – 198.

**Cotes R.** (посм. публ. 1722), Aestimatio errorum in mixta mathesi per variationes partium trianguli. *Opera Misc*. London, 1768, pp. 10 – 68. Перевод в книге Gowing R. (1983), *Roger Cotes – Natural Philosopher.* Cambridge.

**Cubranic R.** (1961), *Geodetski rad R. Boskovica.* Zagreb.

**Czuber E.** (1890), Zur Theorie der Beobachtungsfehler. *Monatsh. Math. Phys.*, Bd. 1, pp. 457 – 464.

--- (1891a), *Theorie der Beobachtungsfehler.* Leipzig.

--- (1891b), Zur Kritik einer Gauss'schen Formel. *Monatsh. Math. Phys*., Bd. 2, pp. 459 – 464.

--- (1903), Über ein Satz der Fehlertheorie und seine Anwendung. *Jahresber. Deutschen Mathematiker-Vereinigung*, Bd. 12, pp. 23 – 30.

**David H. A.** (1957), Some notes on the statistical papers of F. R. Helmert. *Bull. Stat. Soc. New South Wales*, No. 19, pp. 25 – 28. Перепечатка из того же журнала (1954) с комментарием редактора.

**David F. N., Neyman J.** (1938), Extension of the Markoff theorem on least squares. *Stat. Res. Mem.*, vol. 2, pp. 105 – 117.

**Dedekind R.** (1860), Über die Bestimmung der Präzision einer Beobachtungsmethode etc. *Ges. math. Werke*, Bd. 1. Braunschweig, 1930, pp. 95 – 100.

**Delambre J. B. J.** (1827), *Histoire de l'astronomie du dix-huitième siècle.* Paris.

**De Moivre A.** (1712), De mensura sortis. Перевод: *Intern. Stat. Rev.*, vol. 52, 1984, pp. 229 – 262.

--- (1730), *Miscellanea analytica*. London.

--- (1743), *Annuities on Lives.* London. Посм. публ.: De Moivre (1756, pp. 261 – 328).

--- (посм. публ. 1756), *Doctrine of Chances.* London. Прежние издания: 1718, 1738.

**De Morgan A.** (1864), On the theory of errors of observation. *Trans. Cambr. Phil. Soc.*, vol. 10, pp. 409 – 427.

**Descartes R.** (1637), *Discourse de la méthode. Œuvres*, t. 6. Paris, 1982, pp. 1 – 78.

**Dorsey N. E., Eisenhart C.** (1969), On absolute measurements. В книге Ku (1969, pp. 49 – 55).
**Dutka J.** (1990), Adrain and the method of least squares. AHES, vol. 41, pp. 171 – 184.
**Eddington E. S.** (1933), Notes on the method of least squares. *Proc. Phys. Soc.*, vol. 45, pp. 271 – 287.
**Edgeworth F. Y.** (1883), The method of least squares. *Phil. Mag.*, ser. 5, vol. 16, pp. 360 – 375. Также в собр. соч. автора *Writings in Probability, Statistics and Economics*, vol. 2. Cheltenham, UK – Brookfield, US, 1996, pp. 1 – 16.
**Eisenhart C.** (1961), Boscovich and the combination of observations. Перепечатка в книге Pearson E. S., Kendall M. G. (1970, pp. 88 – 100).
--- (1963), Realistic evaluation of the precision and accuracy of instrument calibration systems. Перепечатка в книге Ku (1969, pp. 21 – 47).
--- (1964), The meaning of “least” in least squares. *J. Wash. Acad. Sci.*, vol. 54, pp. 24 – 33. Перепечатка там же, с. 265 – 274.
--- (1976), [Discussion of invited papers on history of statistics]. *Bull. Intern. Stat. Inst.*, vol. 46, pp. 355 – 357.
--- (1978), Gauss. В книге Kruskal, Tanur (1978, vol. 1, pp. 378 – 386).
--- (1983), Laws of error. *Enc. Stat. Sciences*, vol. 6. Редактор Kotz S. и др., 2006, pp. 4052 – 4086.
**Elashoff J. D., Elashoff R. M.** (1978), Effects of errors in statistical assumptions. В книге Kruskal, Tanur (1978, vol. 1, pp. 229 – 250).
**Estienne J. E.** (1926 – 1927), Introduction a une théorie rationelle des erreurs d’observation. *Rev. d’artill.*, t. 97, pp. 421 – 441; t. 98, pp. 542 – 562, t. 100, pp. 471 – 487.
**Euler L.** (1749), Recherches sur la question des inégalités du mouvement de Saturne et de Jupiter. *Opera omnia*, ser. 2, t. 25. Zürich, 1960, pp. 45 – 157.
--- (1778, латинск.), Observationes … Observations on the foregoing dissertation of [Daniel] Bernoulli. *Biometrika*, 1961, vol. 48, pp. 1 – 18. Перепечатка в книге Pearson E. S., Kendall (1970, pp. 167 – 172).
**Farebrother R. W.** (1999), *Fitting Linear Relationships. History of the Calculus of Observations, 1750 – 1900*. New York.
**Fischer P.** (1845), *Lehrbuch der höheren Geodäsie*. Darmstadt.
**Fisher R. A.** (1920), A mathematical examination of the methods of determining the accuracy of an observation. *Monthly Notices Roy. Astron. Soc.*, vol. 80, pp. 758 – 770.
--- (1922), On the mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc.*, vol. A222, pp. 309 – 368.
--- (1925), *Statistical Methods for Research Workers*. В книге автора с отдельной пагинацией (1990), *Stat. Methods, Experimental Design and Scient. Inference*. Oxford.
--- (1939), “Student”. *Annals Eug.*, vol. 9, pp. 1 – 9.
--- (1951), Statistics. В книге *Scient. Thought in the 20th Century*. Редактор A. E. Heath. London, pp. 31 – 55.
**Forbes E. G.** (1974), Mayer. *Dict. Scient. Biogr.*, vol. 9, pp. 232 – 235.

**Forsythe G. E.** (1951), Gauss to Gerling on relaxation. *Math. Tables and Other Aids to Comp.*, vol. 5, No. 36, pp. 255 – 258.

**Fourier J. B. J.** (1826), Sur les résultats moyens. *Œuvres*, t. 2. Paris, 1890, pp. 525 – 545.

**Freudenthal H., Steiner H.-G.** (1966), Aus der Geschichte der Wahrscheinlichkeitstheorie und der mathematische Statistik. В книге *Grundzüge der Mathematik*, Bd. 4. Редактор H. Behnke и др. Göttingen, pp. 149 – 195.

**Friedrich K.** (1937), Allgemeine … Lösung für die Aufgabe der kleinsten Absolutsummen. ZfV, Bd. 66, pp. 305 – 320, 337 – 358.

**Galileo G.** (посм. публ. 1718, итал.), [Thoughts about dice games]. В книге David F. N. *Games, Gods and Gambling*. London, pp. 192 – 195.

--- (1613, итал.), History and demonstrations concerning sunspots. В книге автора *Discoveries and Opinions of Galileo*. Garden City, N. Y., 1957, pp. 88 – 144.

**Gauss C. F.** (посм. публ. 1845), Anwendung der Wahrscheinlichkeitsrechnung auf die Bestimmung der Bilanz für Witwenkassen [часть 2]. *Werke*, Bd. 4. Göttingen, 1880, pp. 125 – 157.

--- (1855), *Méthode des moindres carrés*. Paris.

--- (1870 – 1929), *Werke*, Bde 1 – 12. Göttingen.

--- (1880; 1899; 1927; 1900 – 1909; 1860 – 1865), Переписка с Бесселем, Больяй, Герлингом, Ольберсом и Шумахером. Werke, Ergänzungsreihe, Bde 1 – 5. Hildesheim, 1975, 1987, 1975, 1976, 1975.

--- (1887), *Abhandlungen zur Methode der kleinsten Quadrate*. Редакторы A. Börsch, P. Simon. Vaduz, 1998.

**Gerardy T.** (1977), Die Anfänge von Gauss' geodätische Tätigkeit. ZfV, Bd. 102, pp. 1 – 20.

**Gerling Ch. L.** (1839), *Beiträge zur Geographie Kurhessens.* Cassel.

**Gingerich O.** (1983), Ptolemy, Copernicus and Kepler. В книге *The Great Ideas Today*. Редактор M. J. Adler и др. Chicago, pp. 137 – 180.

**Gini C.** (1946), Gedanken von Theorem von Bernoulli. *Z. f. Volkswirtschaft u. Statistik*, 82. Jg., pp. 401 – 413.

**Glaisher J. W. L.** (1872), On the law of facility of errors of observation. *Mem. Roy. Astron. Soc.*, vol. 39, pp. 75 – 124.

**Gleinsvik P.** (1967), The generalization of the theorem of Jacobi. *Bull. Géo*d., t. 85, pp. 269 – 281.

**Gower B.** (1993), Boscovich on probabilistic reasoning and the combination of observations. В книге *R. J. Boscovich. Vita e attività scient.* Редактор P. Bursill-Hall. Roma, pp. 263 – 279.

**Grafarend E., Harland P.** (1973), *Optimale Design geodätischer Netze.* Deutsche geod. Komm. Bayer. Akad. Wiss., Bd. A74. München.

**Hagen G.** (1837), *Grundzüge der Wahrscheinlichkeitsrechnung*. Berlin. Послед. изд. 1867, 1882.

**Hald A.** (1960), *Statistical Theory with Engineering Applications*. New York – London. Четвертое издание.

--- (1990), *History of Probability and Statistics and Their Applications before 1750*. New York.

--- (1998), *History of Mathematical Statistics from 1750 to 1930*. New York.

**Harter H. L.** (1977, дата предисловия), *Chronological Annotated Bibliography on Order Statistics*, vol. 1. Без места.

**Hauber C. Fr.** (1830 – 1832), Theorie der mittleren Werthe. *Z. Phys. Math.*, Bd. 8, pp. 25 – 56, 147 – 179, 295 – 316; Bd. 9, pp. 302 – 322; Bd. 10, pp. 425 – 457.

**Helmert F. R.** (1868), Studien über rationelle Vermessungen im Gebiete der höhern Geodäsie. *Z. Math. Phys.*, Bd. 13, pp. 73 – 120, 163 – 186. Отдельная публ.: Leipzig, 1868.

--- (1872), *Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate*. Leipzig. Посл. издания: 1907, 1924.

--- (1875a), Über die Berechnung der wahrscheinlichen Fehlers. *Z. Math. Phys.*, Bd. 20, pp. 300 – 303.

--- (1875b), Über die Formeln für den Durchschnittsfehlers. AN, Bd. 85, pp. 353 – 366.

--- (1875c), Discussion der Beobachtungsfehler in Koppe's Vermessung für die Gotthardtunnelachse. ZfV, Bd. 5, pp. 146 – 155.

--- (1876a), Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers. AN, Bd. 88, pp. 113 – 132.

--- (1876b), Über die Wahrscheinlichkeit der Potenzsummen der Beobachtungssfehler. *Z. Math. Phys.*, Bd. 21, pp. 192 – 218.

--- (1877), Über den Maximalfehler einer Beobachtung. ZfV, Bd. 6, pp. 131 – 147.

--- (1886), *Lotabweichungen*, Tl. 1. Berlin.

--- (1904), Zur Ableitung der Formel von Gauss für den mittleren Beobachtungsfehler und ihrer Genauigkeit. *Sitz. Ber. Kgl. Preuss. Akad. Wiss. Berlin*, Hlbbd 1, pp. 950 – 964.

--- (1905), Über die Genauigkeit der Kriterien des Zufalls bei Beobachtungsreihen. Там же, Hlbbd 1, pp. 594 – 612.

**Herschel W.** (1805), On the direction and the motion of the Sun. *Scient. Papers*, vol. 2. London, 1912, pp. 317 – 331.

**Heyde C. C., Seneta E.** (1977), *I. J. Bienaymé*. New York.

**Hogan E. R.** (1977), R. Adrain: American mathematician. *Hist. Math.*, vol. 4, pp. 157 – 172.

**Hulme H. R.** (1940), The statistical theory of errors. *Monthly Notices Roy. Astron. Soc.*, vol. 100, pp. 303 – 314.

**Hulme H. R., Symms L. S. T.** (1939), The law of error and the combination of observations. Там же, vol. 99, pp. 642 – 649.

**Humboldt A.** (  нем./1858), *Cosmos*, vol. 4. New York, 1958. Англ. издание.

**Huygens C.** (1699), Переписка. OC, t. 6. La Haye, 1895, pp. 530 – 532, 537.

**Ivory J.** (1826 – 1830), 11 статей в журнале *Phil. Mag*. об уравнивании маятниковых наблюдений и фигуре Земли. Последняя статья:

--- (1830), On the figure of the Earth. *Phil. Mag.*, vol. 7, pp. 412 – 416.

**Jacobi C. G. J.** (1841, латинск.), Über die Bildung und die Eigenschaften der Determinanten. *Ostwalds Klassiker* No. 77. Leipzig, 1896, pp. 3 – 49.

**Jordan W.** (1877), Über den Maximalfehler einer Beobachtung. ZfV, Bd. 6, pp. 35 – 40.

**Joule J. P.** (1849), On the mechanical equivalent of heat. *Phil. Trans. Roy. Soc.* за 1850, pp. 61 – 82.

**Kac M.** (1939), On a characterization of the normal distribution. В книге автора *Probability, Number Theory and Stat. Physics.* Cambridge, Mass., 1979, pp. 77 – 79.

**Kapteyn J. C.** (1912), Definition of the correlation coefficient. *Monthly Notices Roy. Astron. Soc.*, vol. 72, pp. 518 – 525.

**Kendall M. G.** (1971), The work of E. Abbe. *Biometrika*, vol. 58, pp. pp. 369 – 373. Перепечатка в книге Kendall, Plackett (1977, pp. 331 – 335).

**Kendall M. G., Plackett R. L.**, редакторы (1970), *Studies in the History of Statistics and Probability*. London. Сборник перепечаток.

**Kepler J.** (1606, латинск.), *Über den Neuen Stern im Fuß des Schlangenträger.* Würzburg, 2006.

--- (1609, латинск.), *New Astronomy*. Cambridge, 1992.

**Keynes J. M.** (1921), *Treatise on Probability*. *Coll. Writings*, vol. 8. London, 1973.

**Knobloch E.** (1990), Zur Geneze der Fehlertheorie. В книге *Mathesis rationis. Festschrift für Heinrich Schepers*. Münster, pp. 301 – 327.

**Kohlrausch F.** (1870), *Leitfaden der praktischen Physik*. Leipzig.

**Kries J. von** (1886), *Die Principien der Wahrsxcheinlichkeitsrechnung*. Tübingen, 1927.

**Kruskal W. H.** (1946), Helmert's distribution. *Amer. Math. Monthly*, vol. 53, pp. 435 – 438.

--- (1960), Some remarks on wild observations. В книге Ku (1969, pp. 346 – 348).

--- (1978), Formulas, numbers, words: statistics in prose. В книге *New Directions for Methodology of Soc. and Behavioral Sci.*, No. 9. Редактор D. Fiske. San Francisco, 1981, pp. 93 – 102.

**Kruskal W. H., Tanur J. M.,** редакторы (1978), *International Encyclopedia of Statistics*, vols 1 – 2. New York.

**Ku H. H.** (1967), Statistical concepts in metrology. Перепечатка в книге Ku (1969, pp. 296 – 330).

---, редактор (1969), *Precision Measurement and Calibration. Sel. Nat. Bureau Standards Papers Stat. Concepts and Procedures*. NBS Sp. Publ. 300, vol. 1. Washington.

**Kummel C. H.** (1882), On the composition of errors from single causes of error. AN, Bd. 103, pp. 177 – 206.

**Lagrange J. L.** (1776), Sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations. OC, t. 2. Paris, 1868, pp. 173 – 234.

**Lambert J. H.** (1760), *Photometria*. Augsburg.

--- (1765a), Anmerkungen und Zusätze zur practischen Geometrie. В книге автора *Beyträge zum Gebrauche der Mathematik und deren Anwendung*, Tl. 1. Berlin, pp. 1 – 313.

--- (1765b), Theorie der Zuverlässigkeit der Beobachtungen und Versuche. Там же, pp. 424 – 488.

**Lamont J.** (прим. 1839), Nachricht über meteorologische Bestimmung des Königreiches Bayern. В книге автора *Jahrb. Kgl. Sternwarte bei München* за 1839, pp. 256 – 264, 247 – 249. Пагинация последних трех страниц ошибочна.

--- (1867), Über die Bedeutung arithmetischer Mittelwerthe in der Meteorologie. *Z. Öster. Ges. Met.*, Bd. 2, No. 11, pp. 241 – 247.

**Lancaster H. O.** (1966), Forerunners of the Pearson's chi square. *Austr. J. Stat.*, vol. 8, pp. 117 – 126.
**Laplace P. S.** (1774), Sur la probabilité des causes par les événements. OC, t. 8. Paris, 1891, pp. 27 – 65.
--- (1776), Recherches sur l'intégration des équations différentielles aux différences finies. Там же, pp. 69 – 197.
--- (1781), Sur les probabilités. OC, t. 9. Paris, 1893, pp. 383 – 485.
--- (1784), Sur la chaleur. OC, t. 10. Paris, 1894, pp. 149 – 200.
--- (1792), Sur quelques points du système du monde. OC, t. 11. Paris, 1895, pp. 477 – 558.
--- (1795), Leçons de mathématiques données à l'Ecole normale (1812). OC, t. 14. Paris, 1912, pp. 10 – 177.
--- (1796), *Exposition du système du monde.* Перепечатка издания 1835 г.: OC, t. 6. Paris, 1884.
--- (1798; 1798, год 11 ≈ 1803, 1805, 1825), *Traité de mécanique céleste*. OC, tt. 1 – 5. Paris, 1878, 1878, 1878, 1880, 1882. Перевод на англ. (N. Bowditch): 1832, 1966 (4 тома).
--- (1810a), Sur les approximations des formules etc. OC, t. 12. Paris, 1898, pp. 301 – 345.
--- (1810b), Дополнение к (1810a). Там же, pp. 349 – 353.
--- (1811), Sur les intégrales définies. Там же, pp. 357 – 412.
--- (1812), *Théorie analytique des probabilités*, кн. 2. OC, t. 7, No. 2. Paris, 1886, pp. 181 – 496.
--- (1816), Suppl. 1. Там же, pp. 497 – 530.
--- (1818), Suppl. 2. Там же, pp. 531 – 580.
--- (прим. 1819), Suppl. 3. Там же, pp. 581 – 616.
--- (1821), Sur la détermination des orbites des comètes. OC, t. 13. Paris, 1904, pp. 265 – 267.
--- (1827), Sur le flux et reflux lunaire atmosphérique. Там же, pp. 342 – 358.
**Legendre A. M.** (1805), *Nouvelles méthodes pour la détermination des orbites des comètes*. Paris.
--- (1814), Méthode des moindres quarrés, pour trouver le milieu les plus probable entre les résultats de différentes observations. *Mém.* Cl. sci. math. et phys. *Acad. Sci. Paris*, t. 11, pt. 2 за 1810, pp. 149 – 154.
--- (1820), *Nouvelles méthodes* etc, Supplément 2. Paris.
**Lehmann-Filhés R.** (1887), Über abnorme Fehlervertheilung. AN, Bd. 117, pp. 121 – 132.
**Lévy P.** (1924), La loi de Gauss et les lois exceptionelles. *Bull. Soc. Math. France*, No. 52, pp. 49 – 85.
--- (1925), *Calcul des probabilités*. Paris.
--- (1929), Sur quelques travaux relatifs a la théorie des erreurs. *Bull. sci. math.*, sér. 2, t. 53, pp. 11 – 32.
--- (1970), *Quelques aspects de la pensé d'un mathématicien*. Paris.
**Liagre J. B. J.** (1852), *Calcul des probabilités et théorie des erreurs*. Bruxelles, 1879.
**Lipschitz R.** (1890), Sur la combinaison des observations. *C. r. Acad. Sci. Paris*, t. 111, pp. 163 – 166.
**Maire** [**C.**], **Boscovich** [**R.**] (1770), *Voyage astronomique et géographique dans l'Etat de l'Eglise.* Paris.

**Maupertuis P. L. M.** (1738), Relation du voyage … au cercle polaire pour déterminer la figure de la Terre. *Œuvres*, t. 3. Lyon, 1756, pp. 68 – 175.

--- (1756), Opérations pour déterminer la figure de la Terre et les variations de la pesanteur. *Œuvres*, t. 4. Lyon, 1756, pp. 285 – 346.

**May K. O.** (1972), Gauss. *Dict. Scient. Biogr.*, vol. 5, pp. 298 – 315.

**Mayer** [**T.**] (1750), Abhandlung über die Umwälzung des Mondes um seine Axe. *Kosmogr. Nachr. u. Samml.* за 1748, pp. 52 – 183.

**Meadowcroft L. V.** (1920), On Laplace's theorem on simultaneous errors. *Messenger Math.*, vol. 50, pp. 40 – 48.

**Mendoza E.** (1991), Physics, chemistry and the theory of errors. *Arch. Intern. Hist. Sci.*, vol. 41, pp. 282 – 306.

**Merian P.** (1830), D. Huber. В книге *Verh. Allg. schweiz. Ges. ges. Naturwiss. in ihrer 16. Jahresversamml. zu St. Gallen 1830*. St. Gallen, pp. 145 – 152.

**Merriman M.** (1877), List of writings relating to the method of least squares. Перепечатка в книге Stigler (1980, vol. 1).

**Meyer, Hugo** (1891), *Anleitung zur Bearbeitung meteorologischer Beobachtungen*. Berlin.

**Mises R. von** (1919), Fundamentalsätze der Wahrscheinkeitsrechnung. *Math. Z.*, Bd. 4, pp. 1 – 97. Частичная перепечатка в книге автора *Sel. Papers*, vol. 2. Providence, Rhode Island, 1964, pp. 35 – 56.

**Needham J.** (1962), *Science and Civilisation in China*, vol. 4, pt. 1. Cambridge.

**Neugebauer O.** (1948), Mathematical methods in ancient astronomy. В книге автора *Astronomy and History. Sel. Essays*. New York, 1983, pp. 99 – 127.

--- (1950), The alleged Babylonian discovery of the precession of the equinoxes. Там же, pp. 247 – 254.

--- (1975), *History of Ancient Mathematical Astronomy*, pts 1 – 3. Berlin.

**Neumann J. von и др.** (1941), The mean square successive differences. *Annals Math. Stat.*, vol. 12, pp. 153 – 162.

**Newcomb S.** (1867), *Investigation of the Orbit of Neptune. Smithsonian Contr. to Knowledge*, vol. 15. Отдельная пагинация.

--- (1872), *On the Right Ascensions of the Equatorial Fundamental Stars*. Washington.

--- (1874), *Investigation of the Orbit of Uranus. Smithsonian Contr. to Knowledge*, vol. 19. Отдельная пагинация.

--- (1878), *Researches of the Motion of the Moon*, pt. 1. *Wash. Observations* за 1875, Appendix 2.

--- (1882), Discussions and results of observations. *Astron. Papers Amer. Ephemeris*, vol. 1, pp. 363 – 487.

--- (1886), A generalized theory of the combination of observations. *Amer. J. Math.*, vol. 8, pp. 343 – 366. Перепечатка в книге Stigler (1980, vol. 2).

--- (1892), On the law and the period of the variation of terrestrial latitudes. AN, Bd. 130, pp. 1 – 6.

--- (1895), *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy*. Washington.

--- (1896a), On the variation of the personal equation with the magnitude of the star observed. *Astron. J.*, vol. 16, pp. 65 – 67.

--- (1896b), On the solar motion as a gauge of the stellar distances. Там же, т. 17, с. 41 – 44.

--- (1897a), A new determination of the precessional constant. *Astron. Papers*, vol. 8, pp. 1 – 76.

--- (1897b), То же название, *Astron. J.*, vol. 17, pp. 161 – 167.

--- (1897c), *Problems of Astronomy*. Без места, Univ. Pennsylvania.

--- (1901), On the period of Solar spots. *Astrophys. J.*, vol. 13, pp. 1 – 14.

**Newcomb S., Holden E. S.** (1874), On the possible periodic changes of the Sun's apparent diameter. *Amer. J. Sci.*, ser. 3, vol. 8 (108), pp. 268 – 277.

**Newton R. R.** (1977), *The Crime of Claudius Ptolemy*. Baltimore.

**Neyman J.** (1934), On the two different aspects of the representative method. *J. Roy. Stat. Soc.*, vol. 97, pp. 558 – 625.

**Ogorodnikoff K.** (1928), A method for combining observations. *Астрон. Ж.*, т. 5, № 1, с. 1 – 21.

--- (1929a), On the occurrence of discordant observations. *Monthly Notices Roy. Astron. Soc.*, vol. 88, pp. 523 – 532.

--- (1929b), On a general method of treating observations. *Астрон. Ж.*, т. 6, с. 226 – 244.

**Paucker M. G.** (1819), *Über die Anwendung der Methode der kleinsten Quadratsumme auf physikalische Beobachtungen. Program zur Eröffnung des Lehrkursus auf dem Gymnasium illustre zu Mitau für 1819*. Mitau.

**Pearson E. S., Kendall M. G.**, редакторы (1970), *Studies in the History of Statistics and Probability*. London. Сборник перепечаток.

**Pearson K.** (1894), On the dissection of asymmetrical frequency curves. *Phil. Trans. Roy. Soc.*, vol. A185, pt. 1, pp. 71 – 110.

--- (1898), Cloudiness. *Proc. Roy. Soc.*, vol. 62, pp. 287 – 290.

--- (1900), On a criterion etc. *Phil. Mag*., vol. 50, pp. 157 – 175. Перепечатка в книге автора *Early Statistical Papers.* Cambridge, 1956, pp. 339 – 357.

--- (1920), Notes on the history of correlation. *Biometrika*, vol. 13, pp. 25 – 45. Перепечатка в книге Pearson E. S., Kendall (1970, pp. 185 – 205).

--- (1930), *Life, Letters and Labours of Fr. Galton*, vol. 3A. Cambridge.

--- (1931), Historical note on the distribution of the standard deviation. *Biometrika*, vol. 23, pp. 416 – 418.

--- (1978), *The History of Statistics in the 17th and 18th Centuries*. Редактор E. S. Pearson. London.

**Peters C. A. F.** (1856), Über die Bestimmung des wahrscheinlichen Fehlers. AN, Bd. 44, pp. 29 – 32.

**Pitman E. J. G.** (1939), Estimation of the location and scale parameters. *Biometrika*, vol. 30, pp. 391 – 421.

**Plackett R. L.** (1949), Historical note on the method of least squares. *Biometrika*, vol. 36, pp. 458 – 460.

--- (1950), Some theorems in least squares. Там же, т. 37, с. 149 – 157.

--- (1958), The principle of the arithmetic mean. Там же, т. 45, с. 130 – 135. Перепечатка в книге Pearson E. S., Kendall (1970, pp. 121 – 126).

--- (1972), Discovery of the method of least squares. Там же, т. 59, с. 239 – 251. Перепечатка в книге Kendall, Plackett (1977, pp. 279 – 291).

**Poincaré H.** (опубл. посм. 1980, написано 1921), Résumé analytique [своих собств. работ]. В книге *Math. Heritage of H. Poincaré. Proc. Symp. Indiana Univ. 1980*. Редактор F. E. Browder. Providence, Rhode Island, pp. 257 – 357.

**Poisson S. D.** (1837), *Recherches sur la probabilité des jugements*. Paris, 2003.

**Price D. J.** (1955), Medieval land surveying and topographical maps. *Geogr. J.*, vol. 121, pt. 1, pp. 1 – 10.

**Proverbio E.** (1988), Boscovich's determination of instrumental errors in observation. AHES, vol. 38, pp. 135 – 152.

**Ptolemy** (1956), *Tetrabiblos*. London. Греч. и англ.

--- (1984, англ.), *Almagest*. London.

**Pukelsheim F.** (1993), *Optimal Design of Experiments*. New York.

**Quetelet A.** (1836), Sur l'homme, tt. 1 – 2. Bruxelles.

--- (1846), *Lettres … sur la théorie des probabilités*. Bruxelles.

--- (1849), *Sur le climat de la Belgique*, t. 1. Bruxelles.

--- (1853), *Théorie des probabilités*. Bruxelles.

**Rabinovich N. L.** (1974), Early antecedents of error theory. AHES, vol. 13, pp. 348 – 358.

**Radelet-De Grave P., Scheuber V.** (1979), *Correspondance entre D. Bernoulli et J.-H. Lambert*. Paris.

**Rigaud S. P.** (1832), *Miscellaneous Works and Correspondence of the Rev. James Bradley*. New York, 1972.

**Rosenberger O. A.** (1828), Über die … vorgenommene Gradmessung. AN, Bd. 6, pp. 1 – 32.

**Sabine E.** (1821), An account of experiments to determine the acceleration of the pendulum. *Trans. Roy. Soc.*, pp. 163 – 190.

**Sampson R. A.** (1913), On the law of distribution of errors. *Proc. Fifth Intern. Congr. Mathematicians*. Cambridge, vol. 2, pp. 163 – 173.

**Schmeidler F.** (1984), *Leben und Werk des Königsberger Astronomen F. W. Bessel.* Kelkheim/T.

**Schneider I.** (1975), R. Clausius' Beitrag zur Einführung wahrscheinlichkeitstheoretischer Methoden in die Physik etc. AHES, vol. 14, pp. 237 – 261.

---, редактор (1988), *Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfangen bis 1933*. Darmstadt. Хрестоматия. Почти исключительно на немецком языке.

**Schreiber** [**O.**] (1879), Richtungsbeobachtungen und Winkelbeobachtungen. ZfV, Bd. 8, pp. 97 – 149.

--- (1882), Die Anordnung der Winkelbeobachtungen in Göttinger Basisnetz. ZfV, Bd. 11, pp. 129 – 161.

**Schwerd F. M.** (1822), *Die kleine Speyerer Basis*. Speyer.

**Seal H. L.** (1949), The historical development of the use of generating functions in probability theory. *Bull. Assoc. Actuaires Suisses*, t. 49, pp. 209 – 228. Перепечатка в книге Kendall, Plackett (1977, pp. 67 – 86).

--- (1967), The historical development of the Gauss linear model. *Biometrika*, vol. 54, pp. 1 – 24. Перепечатка в книге Pearson E. S., Kendall (1970, pp. 207 – 230).

**Seeliger H.** (1900), Über die Verteilung der nach einer Ausgleichung übrig bleibenden Fehler. *Sitz. Ber.* math. phys. Kl. *Kgl. Bayer. Akad. Wiss. München*, Bd. 29 за 1899, pp. 3 – 21.

**Seidel L.** (1863), Über eine Anwendung der Wahrscheinlichkeits-rechnung. *Sitz. Ber.* math. phys. Kl. *Bayer. Akad. Wiss. München*, Bd. 2 за 1863, pp. 320 – 350.

--- (1865), Über den … Zusammenhang … zwischen der Häufigkeit der Typhus-Erkrankungen und dem Stande des Grundwassers. *Z. Biol.*, Bd. 1, pp. 221 – 236.

--- (1866), Vergleichung der Schwankungen der Regenmengen mit der Schwankungen in der Häufigkeit des Typhus. Там же, т. 2, с. 145 – 177.

**Sheynin O. B.** (1963), Adjustment of a trilateration figure by frame structure analogue. *Empire Surv. Rev.*, vol. 17, No. 127, pp. 55 – 56.

--- (1966), Origin of the theory of errors. *Nature*, vol. 211, pp. 1003 – 1004.

--- (1971a), Newton and the classical theory of probability. AHES, vol. 7, pp. 217 – 243.

--- (1971b), Lambert's work in probability. Там же, pp. 244 – 256.

--- (1972a), On the mathematical treatment of observations by Euler. AHES, vol. 9, pp. 45 – 56.

--- (1972b), D. Bernoulli's work on probability. Перепечатка в книге Kendall, Plackett (1977, pp. 105 – 132).

--- (1973a), Finite random sums. AHES, vol. 9, pp. 275 – 305.

--- (1973b), Boscovich's work on probability. Там же, pp. 306 – 324.

--- (1973c), Mathematical treatment of observations. AHES, vol. 11, pp. 97 – 126.

--- (1975a), Kepler as a statistician. *Bull. Intern. Stat. Inst.*, vol. 46, pp. 341 – 354.

--- (1977), Laplace's theory of errors. AHES, vol. 17, pp. 1 – 61.

--- (1979), Gauss and the theory of errors. AHES, vol. 20, pp. 21 – 72.

--- (1982), On the history of medical statistics. AHES, vol. 26, pp. 241 – 286.

--- (1984a), On the history of the statistical method in astronomy. AHES, vol. 29, pp. 151 – 199.

--- (1984b), On the history of the statistical method in meteorology. AHES, vol. 31, pp. 53 – 95.

--- (1986), Quetelet as a statistician. AHES, vol. 36, pp. 281 – 325.

--- (1989), Markov's work on probability. AHES, vol. 39, pp. 337 – 377; vol. 40, p. 387.

--- (1992a), Al-Biruni and the mathematical treatment of observations. *Arabic Sciences and Phil.*, vol. 2, pp. 299 – 306.

--- (1993a), On the history of the principle of least squares. AHES, vol. 46, pp. 39 – 54.

--- (1993b), Treatment of observations in early astronomy. Там же, pp. 153 – 192.

--- (1994a), Theory of errors. В книге *Companion Enc. Hist. Phil. Math. Sciences*, vol. 2. Редактор I. Grattan-Guinness. London, pp. 1315 – 1324.

--- (1994b), Gauss and geodetic observations. AHES, vol. 46, pp. 253 – 283.

--- (1994c), Ivory's treatment of pendulum observations. *Hist. Math.*, vol. 21, pp. 174 – 184.
--- (1994d), Chebyshev's lectures on the theory of probability. AHES, vol. 46, pp. 321 – 340.
--- (1994e), Bertrand's work on probability. AHES, vol. 48, pp. 155 – 199.
--- (1995a), Helmert's work in the theory of errors. AHES, vol. 49, pp. 73 – 104.
--- (1995b), Density curves in the theory of errors. Там же, pp. 163 – 196.
--- (1996a), *History of the Theory of Errors*. Egelsbach, Hänsel-Hohenhausen.
--- (1996b), Selection and treatment of observations by Mendeleev. *Hist. Math.*, vol. 23, pp. 54 – 67.
--- (1999a), The discovery of the principle of least squares. *Historia Scientiarum*, vol. 8, pp. 249 – 264.
--- (1999b), Gauss and the method of least squares. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 219, pp. 458 – 467.
--- (1999c), Statistics, definitions of. *Enc. Stat. Sciences*, vol. 12, 2006. Редактор S. Kotz и др., pp. 8128 – 8135.
--- (2000b), Bessel: some remarks on his work. *Historia Scientiarum*, vol. 10, pp. 77 – 83.
--- (2001a), Pirogov as a statistician. Там же, pp. 213 – 225.
--- (2001b), Gauss. В книге Heyde C.C., Seneta E., редакторы (2001), *Statisticians of the Centuries*. New York, pp. 119 – 122.
--- (2001c), Gauss, Bessel and the adjustment of triangulation. *Historia Scientiarum*, vol. 11, pp. 168 – 175.
--- (2002a), Newcomb as a statistician. Там же, т. 12, с. 142 – 167.
--- (2003), Nekrasov's work on probability: the background. AHES, vol. 57, pp. 337 – 353.
--- (2005b), Early discovery of sunspots. *Journal*, vol. 99, No. 3, p. 83.
--- (2006b), Markov's work on the treatment of observations. *Historia scientiarum*, vol. 16, pp. 80 – 95.
**Shoesmith E.** (1985), T. Simpson and the arithmetic mean. *Hist. Math.*, vol. 12, pp. 352 – 355.
**Short J.** (1763), Second paper concerning the parallax of the Sun. *Phil. Trans. Roy. Soc.*, vol. 53, pp. 300 – 342.
**Simpson J. Y.** (1869 – 1870), Hospitalism. *Works*, vol. 2. Edinburgh, 1871, pp. 289 – 405.
**Simpson T.** (1740), *The Nature and the Laws of Chance*. London.
--- (1756), A letter … on the advantage of taking the mean etc. *Phil. Trans. Roy. Soc.*, vol. 49, pt. 1 за 1755, pp. 82 – 93.
--- (1757), То же название, расширенный вариант. В книге автора *Misc. Notes on Some Curious Subjects* etc. London, pp. 64 – 75. Перевод обоих мемуаров см. Шейнин (2006а, с. 116 – 129 .
--- (посм. публ. 1775), *The Docrine of Annuities and Reversions*. London.
**Spieß W.** (1939), Kann man für D. Huber Ansprüche als Erfinder der "Methode der kleinste Quadrate" geltend machen? *Schweiz. Z. Vermessungswesen u. Kulturtechnik*, Bd. 37, pp. 11 – 17, 21 – 23.

**Sprott D. A.** (1978), Gauss’ contribution to statistics. *Hist. Math.*, vol. 5, pp. 183 – 203.

**Stigler S. M.** (1973a), Newcomb, Daniell and the history of robust estimation, 1885 – 1920. *J. Amer. Stat. Assoc.*, vol. 68, pp. 872 – 879. Перепечатка в книге Kendall, Plackett (1977, pp. 410 – 417).

--- (1973b), Laplace, Fisher and the discovery of the concept of sufficiency. *Biometrika*, vol. 60, pp. 439 – 445. Перепечатка: там же, с. 271 – 277.

--- (1978), Mathematical statistics in the early States. *Ann. Stat.*, vol. 6, pp. 239 – 265.

---, редактор (1980), *American Contributions to Mathematical Statistics in the 19th Century*, vols 1 – 2. New York. Сборник перепечаток. Единая пагинация в томах отсутствует.

--- (1981), Gauss and the invention of least squares. *Ann. Stat.*, vol. 9, pp. 465 – 474.

--- (1984), Boscovich, Simpson and a 1760 note. *Biometrika*, vol. 71, pp. 615 – 620.

--- (1986), *History of Statistics.* Cambridge, Mass.

--- (1999), *Statistics on the Table*. Cambridge, Mass. Сборник перепечаток собственных статей.

**Strasser G.** (1957), *Ellipsoidische Parameter der Erdfigur (1800 – 1950).* München.

**Strecker A.** (1846), Über die Atomgewichte des Silbers und Kohlenstoffs. *Annalen Chem. Pharm.*, t. 59, pp. 265 – 284.

**Tilling L.** (1975), Early experimental graphs. *Brit. J. Hist. Sci.*, vol. 8, pp. 193 – 213.

**Todhunter I.** (1865), *History of the Mathematical Theory of Probability.* New York, 1949, 1965.

--- (1869), On the method of least squares. *Trans. Cambr. Phil. Soc.*, vol. 11, pp. 219 – 238.

**Toomer G. J.** (1974), Hipparchus on the distances of the sun and the moon. AHES, vol. 14, pp. 126 – 142.

**Walker H. M.** (1928), The relation of Plana and Bravais to theory of correlation. *Isis*, vol. 10, pp. 466 – 484.

**Weiling F.** (1975), Mendel sowie die von Pettenkoffer angeregten Untersuchungen des Zusammenhanges von Cholera und Typhus-Massenerkrankungen mit dem Grundwasserstand. *Sudhoffs Arch.*, Bd. 59, pp. 1 – 19.

**Wesley W. G.** (1978), The accuracy of Tycho Brahe’s instruments. *J. Hist. Astron.*, vol. 9, pp. 42 – 53.

**Whittaker E. T., Robinson G.** (1924), *Calculus of Observations*. London, 1949. Много иных изданий.

**Wilson C.** (1984), The sources of Ptolemy’s parameters. *J. Hist. Astron.*, vol. 15, pp. 37 – 47.

**Wolf R.** (1858), Daniel Huber von Basel. В книге *Biographien zur Kulturgeschichte der Schweiz*. 1. Cyclus, pp. 441 – 462.

**Yarochenko S.** (1893b), Sur la méthode des moindres carrés. *Bull. Sci. Math.*, sér. 2, t. 17, pp. 113 – 125.

**Yule G. U., Kendall M. G.** (1937), *Introduction to Theory of Statistics.* London, 1958.

**Zach F. X. von** (1813), Sur le degré du méridien. *Mém. Acad. Imp. Sci., Littérature, Beaux-Arts Turin* за 1811 – 1812. Sci. math. et phys., pp. 81 – 216.
**Zoch R. T.** (1935), On the postulate of the arithmetic mean. *Annals Math. Stat.*, vol. 6, pp. 171 – 182.