

# Обработка результатов учебного эксперимента

П.В. Попов, А.А. Нозик

22 сентября 2018 г.

Данный документ является рабочей версией пособия. Оформление будет изменяться по мере доработки. Также ожидаются небольшие изменения по содержанию.

Любые замечания по содержанию или оформлению следует оставлять в репозитории <https://github.com/altavir/mipt-stat-book>. Также замечания можно отправлять по электронной почте или через телеграм: <http://npm.mipt.ru/about.html>.

Дата последнего обновления: 22 сентября 2018 г.

## ОГЛАВЛЕНИЕ

<b>1</b>	<b>Измерения и погрешности</b>	<b>3</b>
1.1	Результат измерения . . . . .	4
1.2	Многократные измерения . . . . .	6
1.3	Классификация погрешностей . . . . .	8
1.3.1	Случайные погрешности . . . . .	10
1.3.2	Систематические погрешности . . . . .	11
<b>2</b>	<b>Элементы теории ошибок</b>	<b>12</b>
2.1	Случайная величина . . . . .	12
2.2	Нормальное распределение . . . . .	16
2.3	Независимые величины . . . . .	19
2.3.1	Дисперсия суммы независимых величин . . . . .	20
2.4	Погрешность среднего . . . . .	21
2.5	Результирующая погрешность опыта . . . . .	22
2.6	Обработка косвенных измерений . . . . .	24
2.6.1	Случай одной переменной . . . . .	25
2.6.2	Случай многих переменных . . . . .	26
<b>3</b>	<b>Оценка параметров</b>	<b>28</b>
3.1	Метод минимума хи-квадрат . . . . .	29
3.2	Метод максимального правдоподобия. . . . .	30
3.3	Метод наименьших квадратов (МНК). . . . .	31
3.4	Проверка качества аппроксимации . . . . .	31
3.5	Оценка погрешности параметров . . . . .	32
3.6	Методы построения наилучшей прямой . . . . .	33
3.6.1	Метод наименьших квадратов . . . . .	33
3.6.2	Погрешность МНК в линейной модели . . . . .	34

3.6.3	Недостатки и условия применимости МНК . . . . .	35
3.6.4	Метод хи-квадрат построения прямой . . . . .	35
<b>4</b>	<b>Рекомендации по выполнению и представлению результатов работы</b>	<b>36</b>
4.1	Проведение измерений . . . . .	36
4.1.1	Правила ведения лабораторного журнала . . . . .	36
4.1.2	Подготовка к работе . . . . .	38
4.1.3	Начало работы . . . . .	39
4.1.4	Выбор количества измерений . . . . .	39
4.1.5	Измерения . . . . .	40
4.1.6	Расчёты, анализ и представление результатов . . . . .	41
4.2	Анализ инструментальных погрешностей . . . . .	41
4.3	Отчёт о работе . . . . .	43
4.3.1	Требования к содержанию разделов . . . . .	44
4.3.2	Правила округления . . . . .	45
4.3.3	Построение графиков . . . . .	46
4.3.4	Рекомендации по оформлению графиков . . . . .	49
4.4	Некоторые типичные ошибки обработки данных . . . . .	50
<b>5</b>	<b>Приложение</b>	<b>54</b>
5.1	Корреляции . . . . .	54
5.2	Свойства точечных оценок . . . . .	55
5.3	Максимальная эффективность оценки (граница Рао–Крамера)	56
5.4	Погрешности коэффициентов построения прямой . . . . .	57
5.5	Многопараметрические оценки . . . . .	59
5.5.1	Двумерный случай . . . . .	59
5.5.2	Многомерный случай . . . . .	60
5.5.3	Использование пакета <code>scipy</code> для построение оценки . . . . .	61

## ПРЕДИСЛОВИЕ

Данное пособие содержит краткое изложение основных понятий и методов, необходимых для корректной обработки результатов экспериментов в учебной лаборатории, а также указания по представлению результатов лабораторных работ, с тем чтобы оформленные студентами отчёты в целом соответствовали сложившимся современным стандартам оформления научных публикаций.

Умение работать с погрешностями, или «ошибками», является важной частью любого научного эксперимента на всех его этапах. Так, при подготовке и проведении эксперимента необходимо знать точность исполь-

зуемых приборов, уметь находить пути возможного уменьшения ошибок, разумно организовать сами измерения и правильно оценивать точность полученных значений. На этапе обработки возникает необходимость пересчитывать возможную ошибку в конечных результатах по известным оценкам погрешностей в исходных данных. А на самом важном этапе — интерпретации результатов эксперимента — без знания точности проведённых измерений и без корректной статистической обработки невозможно делать обоснованные выводы в пользу той или иной физической модели, той или иной гипотезы.

Пособие может быть рекомендовано студентам первого курса для первичного ознакомления с предметом, либо в качестве «шпаргалки» по основным формулам. Более подробное изложение материала, снабжённое большим количеством практических примеров, можно найти в пособиях [1, 2, 3]. Студентам старших курсов, знакомым с основами теории вероятностей и математической статистики, и желающим на более глубоком уровне познакомиться с современными методиками обработки данных, можно порекомендовать начать с книг [4, 5].

## 1. ИЗМЕРЕНИЯ И ПОГРЕШНОСТИ

Свойства физического объекта (явления, процесса) определяются набором количественных характеристик — *физических величин*. Как правило, результат измерения представляет собой *число*, задающее отношение измеряемой величины к некоторому *эталоноу*. Сравнение с эталоном может быть как *прямым* (проводится непосредственно экспериментатором), так и *косвенным* (проводится с помощью некоторого прибора, которому экспериментатор доверяет). Полученные таким образом величины имеют *размерность*, определяемую выбором эталона.

**Замечание.** Результатом измерения может также служить количество отсчётов некоторого события, логическое утверждение (да/нет) или даже качественная оценка (сильно/слабо/умеренно). Мы ограничимся наиболее типичным для физики случаем, когда результат измерения может быть представлен в виде *числа* или *набора чисел*.

Взаимосвязь между различными физическими величинами может быть описана *физическими законами*, представляющими собой идеализированную *модель* действительности. Конечной целью любого физического эксперимента (в том числе и учебного) является проверка адекватности или уточнение параметров таких моделей.

## 1.1. Результат измерения

Рассмотрим простейший пример: измерение длины стержня с помощью линейки. Линейка проградуирована производителем с помощью некоторого эталона длины — таким образом, сравнивая длину стержня с ценой деления линейки, мы выполняем косвенное сравнение с общепринятым стандартным эталоном.

Допустим, мы приложили линейку к стержню и увидели на шкале некоторый результат  $x = x_{\text{изм}}$ . Можно ли утверждать, что  $x_{\text{изм}}$  — это длина стержня?

Во-первых, значение  $x$  *не может быть задано точно*, хотя бы потому, что оно обязательно *округлено* до некоторой значащей цифры: если линейка «обычная», то у неё есть *цена деления*; а если линейка, к примеру, «лазерная» — у неё высвечивается *конечное число значащих цифр* на дисплее.

Во-вторых, мы никак не можем быть уверенны, что длина стержня *на самом деле* такова хотя бы с точностью до ошибки округления. Действительно, мы могли приложить линейку не вполне ровно; сама линейка могла быть изготовлена не вполне точно; стержень может быть не идеально цилиндрическим и т.п.

И, наконец, если пытаться хотя бы гипотетически переходить к бесконечной точности измерения, *теряет смысл само понятие* «длины стержня». Ведь на масштабах атомов у стержня нет чётких границ, а значит говорить о его геометрических размерах в таком случае крайне затруднительно!

Итак, из нашего примера видно, что никакое физическое измерение не может быть произведено абсолютно точно, то есть **у любого измерения есть погрешность**.

**Замечание.** Также используют эквивалентный термин *ошибка измерения* (от *англ. error*). Подчёркнём, что смысл этого термина отличается от общепотребительного бытового: если физик говорит «в измерении есть ошибка», — это не означает, что оно неправильно и его надо переделать. Имеется ввиду лишь, что это измерение *неточно*, то есть имеет *погрешность*.

Количественно погрешность можно было бы определить как разность между измеренным и «истинным» значением длины стержня:  $\delta x = x_{\text{изм}} - x_{\text{ист}}$ . Однако на практике такое определение использовать нельзя: во-первых, из-за неизбежного наличия погрешностей «истинное» значение измерить невозможно, и во-вторых, само «истинное» значение может отличаться в разных измерениях (например, стержень неровный или изогнутый, его торцы дрожат из-за тепловых флуктуаций и т.д.). Поэтому говорят обычно об *оценке* погрешности.

Об измеренной величине также часто говорят как об *оценке*, подчеркивая, что эта величина не точна и зависит не только от физических свойств исследуемого объекта, но и от процедуры измерения.

**Замечание.** Термин *оценка* имеет и более формальное значение. Оценкой называют результат процедуры получения значения параметра или параметров физической модели, а также иногда саму процедуру. Теория оценок является подразделом математической статистики. Некоторые ее положения изложены в главе 3, но для более серьезного понимания следует обратиться к [5].

Для оценки значения физической величины корректно использовать не просто некоторое фиксированное число  $x_{\text{изм}}$ , а *интервал* (или *диапазон*) значений, в пределах которого может лежать её «истинное» значение. В простейшем случае этот интервал может быть записан как

$$x = x_{\text{изм}} \pm \delta x,$$

где  $\delta x$  — *абсолютная* величина погрешности. Эта запись означает, что исследуемая величина лежит в интервале  $x \in (x_{\text{изм}} - \delta x; x_{\text{изм}} + \delta x)$  с некоторой достаточно большой долей вероятности (более подробно о вероятностном содержании интервалов см. п. 2.2). Для наглядной оценки точности измерения удобно также использовать *относительную* величину погрешности:

$$\varepsilon_x = \frac{\delta x}{x_{\text{изм}}}.$$

Она показывает, насколько погрешность мала по сравнению с самой измеряемой величиной (её также можно выразить в процентах:  $\varepsilon = \frac{\delta x}{x} \cdot 100\%$ ).

**Пример.** Штангенциркуль — прибор для измерения длин с ценой деления 0,1 мм. Пусть диаметр некоторой проволоки равен 0,37 мм. Считая, что абсолютная ошибка составляет половину цены деления прибора, результат измерения можно будет записать как  $d = 0,40 \pm 0,05$  мм (или  $d = (40 \pm 5) \cdot 10^{-5}$  м). Относительная погрешность составляет  $\varepsilon \approx 13\%$ , то есть точность измерения весьма посредственная — поскольку размер объекта близок к пределу точности прибора.

**О необходимости оценки погрешностей.** Измерим длины двух стержней  $x_1$  и  $x_2$  и сравним результаты. Можно ли сказать, что стержни одинаковы или различны?

Казалось бы, достаточно проверить, справедливо ли  $x_1 = x_2$ . Но *никакие* два результата измерения не равны друг другу с абсолютной точностью! Таким образом, без указания погрешности измерения ответ на этот вопрос дать *невозможно*.

С другой стороны, если погрешность  $\delta x$  известна, то можно утверждать, что если измеренные длины одинаковы *в пределах погрешности опыта*, если  $|x_2 - x_1| < \delta x$  (и различны в противоположном случае).

Итак, без знания погрешностей невозможно сравнить между собой никакие два измерения, и, следовательно, невозможно сделать *никаких* значимых выводов по результатам эксперимента: ни о наличии зависимостей между величинами, ни о практической применимости какой-либо теории, и т. п. В связи с этим задача правильной оценки погрешностей является крайне важной, поскольку существенное занижение или завышение значения погрешности (по сравнению с реальной точностью измерений) ведёт к *неправильным выводам*.

В физическом эксперименте (в том числе лабораторном практикуме) оценка погрешностей должна проводиться *всегда* (даже когда составители задания забыли упомянуть об этом).

## 1.2. Многократные измерения

Проведём серию из  $n$  *одинаковых* (*однотипных*) измерений одной и той же физической величины (например, многократно приложим линейку к стержню) и получим ряд значений

$$\mathbf{x} = \{x_1, x_2, \dots, x_n\}.$$

Что можно сказать о данном наборе чисел и о длине стержня? И можно ли увеличивая число измерений улучшить конечный результат?

Если цена деления самой линейки достаточно мала, то как нетрудно убедиться на практике, величины  $\{x_i\}$  почти наверняка окажутся *различными*. Причиной тому могут быть самые разные обстоятельства, например: у нас недостаточно остроты зрения и точности рук, чтобы каждый раз прикладывать линейку одинаково; стенки стержня могут быть слегка неровными; у стержня может и не быть определённой длины, например, если в нём возбуждены звуковые волны, из-за чего его торцы колеблются, и т. д.

В такой ситуации результат измерения интерпретируется как *случайная величина*, описываемая некоторым *вероятностным* законом (*распределением*). Подробнее о случайных величинах и методах работы с ними см. гл. 2.

По набору результатов  $\mathbf{x}$  можно вычислить их среднее арифметическое:

$$\langle x \rangle = \frac{x_1 + x_2 + \dots + x_n}{n} \equiv \frac{1}{n} \sum_{i=1}^n x_i. \quad (1.1)$$

Это значение, вычисленное по результатам конечного числа  $n$  измерений, принято называть *выборочным* средним. Здесь и далее для обозначения выборочных средних будем использовать угловые скобки.

Кроме среднего представляет интерес и то, насколько сильно варьируются результаты от опыта к опыту. Определим отклонение каждого измерения от среднего как

$$\Delta x_i = x_i - \langle x \rangle, \quad i = 1 \dots n.$$

Разброс данных относительно среднего принято характеризовать *средне-квадратичным отклонением*:

$$s = \sqrt{\frac{\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_n^2}{n}} = \sqrt{\frac{1}{n} \sum_{i=1}^n \Delta x_i^2} \quad (1.2)$$

или кратко

$$s^2 = \langle (x - \langle x \rangle)^2 \rangle. \quad (1.3)$$

Значение среднего квадрата отклонения  $s^2$  называют выборочной *дисперсией*.

Будем увеличивать число измерений  $n$  ( $n \rightarrow \infty$ ). Если объект измерения и методика достаточно стабильны, то отклонения от среднего  $\Delta x_i$  будут, во-первых, относительно малы, а во-вторых, положительные и отрицательные отклонения будут встречаться примерно одинаково часто. Тогда при вычислении (1.1) почти все отклонения  $\Delta x_i$  скомпенсируются и можно ожидать, что выборочное среднее при  $n \gg 1$  будет стремиться к некоторому пределу:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$$

Тогда предельное значение  $\bar{x}$  можно отождествить с «истинным» средним для исследуемой величины.

Предельную величину среднеквадратичного отклонения при  $n \rightarrow \infty$  обозначим как

$$\lim_{n \rightarrow \infty} \sqrt{\frac{1}{n} \sum_{i=1}^n \Delta x_i^2} = \sigma.$$

**Замечание.** В общем случае указанные пределы могут и не существовать. Например, если измеряемый параметр меняется во времени или в результате самого измерения, либо испытывает слишком большие случайные скачки и т. п. Такие ситуации требуют особого рассмотрения и мы на них не останавливаемся.



**Замечание.** Если  $n$  мало ( $n < 10$ ), для оценки среднеквадратичного отклонения математическая статистика рекомендует вместо формулы (1.3) использовать *исправленную* формулу (подробнее см. п. 5.2):

$$s_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n \Delta x_i^2, \quad (1.4)$$

где произведена замена  $n \rightarrow n-1$ . Величину  $s_{n-1}$  часто называют *стандартным отклонением*.

Итак, можно по крайней мере надеяться на то, что результаты небольшого числа измерений имеет не слишком большой разброс, так что величина  $\langle x \rangle$  может быть использована как приближенное значение (*оценка*) истинного значения  $\langle x \rangle \approx \bar{x}$ , а увеличение числа измерений позволит уточнить результат.

Многие случайные величины подчиняются так называемому *нормальному закону* распределения (подробнее см. Главу 2). Для таких величин могут быть строго доказаны следующие свойства:

- при многократном повторении эксперимента большая часть измерений ( $\sim 68\%$ ) попадает в интервал  $\bar{x} - \sigma < x < \bar{x} + \sigma$  (см. п. 2.2).
- выборочное среднее значение  $\langle x \rangle$  оказывается с большей вероятностью ближе к истинному значению  $\bar{x}$ , чем каждое из измерений  $\{x_i\}$  в отдельности. При этом ошибка вычисления среднего убывает пропорционально корню из числа опытов  $\sqrt{n}$  (см. п. 2.4).

**Упражнение.** Показать, что

$$s^2 = \langle x^2 \rangle - \langle x \rangle^2. \quad (1.5)$$

то есть дисперсия равна разности среднего значения квадрата  $\langle x^2 \rangle = \frac{1}{n} \sum_{i=1}^n x_i^2$

и квадрата среднего  $\langle x \rangle^2 = \left( \frac{1}{n} \sum_{i=1}^n x_i \right)^2$ .

### 1.3. Классификация погрешностей

Чтобы лучше разобраться в том, нужно ли многократно повторять измерения, и в каком случае это позволит улучшить результаты опыта, проанализируем источники и виды погрешностей.

В первую очередь, многократные измерения позволяют проверить *воспроизводимость* результатов: повторные измерения в *одинаковых* условиях, должны давать близкие результаты. В противном случае исследование будет существенно затруднено, если вообще возможно. Таким образом, многократные измерения *необходимы* для того, чтобы убедиться как

в надёжности методики, так и в существовании измеряемой величины как таковой.

При любых измерениях возможны грубые ошибки — *промахи* (англ. miss). Это «ошибки» в стандартном понимании этого слова — возникающие по вине экспериментатора или в силу других непредвиденных обстоятельств (например, из-за сбоя аппаратуры). Промахов, конечно, нужно избегать, а результаты таких измерений должны быть по возможности исключены из рассмотрения.

Как понять, является ли «аномальный» результат промахом? Вопрос этот весьма непрост. В литературе существуют статистические критерии отбора промахов, которыми мы, однако, настоятельно *не рекомендуем* пользоваться (по крайней мере, без серьёзного понимания последствий такого отбора). Отбрасывание аномальных данных может, во-первых, привести к тенденциозному искажению результата исследований, а во-вторых, так можно упустить открытие неизвестного эффекта. Поэтому при научных исследованиях необходимо максимально тщательно проанализировать причину каждого промаха, в частности, многократно повторив эксперимент. Лишь только если факт и причина промаха установлены вполне достоверно, соответствующий результат можно отбросить.

**Замечание.** Часто причины аномальных отклонений невозможно установить на этапе обработки данных, поскольку часть информации о проведении измерений к этому моменту утеряна. Единственным способ борьбы с этим — это максимально подробное описание всего процесса измерений в *лабораторном журнале*. Подробнее об этом см. п. 4.1.1.

При многократном повторении измерения одной и той же физической величины погрешности могут иметь *систематический* либо *случайный* характер. Назовём погрешность *систематической*, если она повторяется от опыта к опыту, сохраняя свой знак и величину, либо *закономерно* меняется в процессе измерений. *Случайные* (или *статистические*) погрешности меняются хаотично при повторении измерений как по величине, так и по знаку, и в изменениях не прослеживается какой-либо закономерности.

Кроме того, удобно разделять погрешности по их происхождению. Можно выделить

- *инструментальные* (или *приборные*) погрешности, связанные с несовершенством конструкции (неточности, допущенные при изготовлении или вследствие старения), ошибками калибровки или ненормативными условиями эксплуатации измерительных приборов;
- *методические погрешности*, связанные с несовершенством теоретической модели явления (использование приближенных формул и

моделей явления) или с несовершенством методики измерения (например, влиянием взаимодействия прибора и объекта измерения на результат измерения);

- *естественные погрешности*, связанные со случайным характером измеряемой физической величины — они являются не столько «ошибками» измерения, сколько характеризуют природу изучаемого объекта или явления.

**Замечание.** Разделение погрешностей на систематические и случайные не является однозначным и зависит от постановки опыта. Например, производя измерения не одним, а несколькими однотипными приборами, мы переводим систематическую приборную ошибку, связанную с неточностью шкалы и калибровки, в случайную. Разделение по происхождению также условно, поскольку любой прибор подвержен воздействию «естественных» случайных и систематических ошибок (шумы и наводки, тряска, атмосферные условия и т. п.), а в основе работы прибора всегда лежит некоторое физическое явление, описываемое не вполне совершенной теорией.

### 1.3.1. Случайные погрешности

Случайный характер присущ большому количеству различных физических явлений, и в той или иной степени проявляется в работе всех без исключения приборов. Случайные погрешности обнаруживаются просто при многократном повторении опыта — в виде хаотичных изменений (*флуктуаций*) значений  $\{x_i\}$ .

Если случайные отклонения от среднего в большую или меньшую стороны примерно равновероятны, можно рассчитывать, что при вычислении среднего арифметического (1.1) эти отклонения скомпенсируются, и погрешность результирующего значения  $\langle x \rangle$  будет меньше, чем погрешность отдельного измерения.

Случайные погрешности бывают связаны, например,

- *с особенностями используемых приборов*: техническими недостатками (люфт в механических приспособлениях, сухое трение в креплении стрелки прибора), с естественными (тепловой и дробовой шуму в электрических цепях, тепловые флуктуации и колебания измерительных устройств из-за хаотического движения молекул, космическое излучение) или техногенными факторами (тряска, электромагнитные помехи и наводки);
- *с особенностями и несовершенством методики измерения* (ошибка при отсчёте по шкале, ошибка времени реакции при измерениях с секундомером);

- с несовершенством объекта измерений (неровная поверхность, неоднородность состава);
- со случайным характером исследуемого явления (радиоактивный распад, броуновское движение).

Остановимся несколько подробнее на двух последних случаях. Они отличаются тем, что случайный разброс данных в них порождён непосредственно объектом измерения. Если при этом приборные погрешности малы, то «ошибка» эксперимента возникает лишь в тот момент, когда мы по своей воле совершаем замену ряда измеренных значений на некоторое среднее  $\{x_i\} \rightarrow \langle x \rangle$ . Разброс данных при этом характеризует не точность измерения, а сам исследуемый объект или явление. Однако с математической точки зрения приборные и «естественные» погрешности *неразличимы* — глядя на одни только экспериментальные данные невозможно выяснить, что именно явилось причиной их флуктуаций: сам объект исследования или иные, внешние причины. Таким образом, для исследования естественных случайных процессов необходимо сперва отдельно исследовать и оценить случайные инструментальные погрешности и убедиться, что они достаточно малы.

### 1.3.2. Систематические погрешности

Систематические погрешности, в отличие от случайных, невозможно обнаружить, исключить или уменьшить просто многократным повторением измерений. Они могут быть обусловлены, во-первых, неправильной работой приборов (*инструментальная погрешность*), например, сдвигом нуля отсчёта по шкале, деформацией шкалы, неправильной калибровкой, искажениями из-за не нормативных условий эксплуатации, искажениями из-за износа или деформации деталей прибора, изменением параметров прибора во времени из-за нагрева и т.п. Во-вторых, их причиной может быть ошибка в интерпретации результатов (*методическая погрешность*), например, из-за использования слишком идеализированной физической модели явления, которая не учитывает некоторые значимые факторы (так, при взвешивании тел малой плотности в атмосфере необходимо учитывать силу Архимеда; при измерениях в электрических цепях может быть необходим учет неидеальности амперметров и вольтметров и т. д.).

Систематические погрешности условно можно разделить на следующие категории.

1. Известные погрешности, которые могут быть достаточно точно вычислены или измерены. При необходимости они могут быть учтены непосредственно: внесением поправок в расчётные формулы или в результаты измерений. Если они малы, их можно отбросить, чтобы упростить вычисления.

2. Погрешности известной природы, конкретная величина которых неизвестна, но максимальное значение вносимой ошибки может быть оценено теоретически или экспериментально. Такие погрешности неизбежно присутствуют в любом опыте, и задача экспериментатора — свести их к минимуму, совершенствуя методики измерения и выбирая более совершенные приборы.

Чтобы оценить величину систематических погрешностей опыта, необходимо учесть паспортную точность приборов (производитель, как правило, гарантирует, что погрешность прибора не превосходит некоторой величины), проанализировать особенности методики измерения, и по возможности, провести контрольные опыты.

3. Погрешности известной природы, оценка величины которых по каким-либо причинам затруднена (например, сопротивление контактов при подключении электронных приборов). Такие погрешности должны быть обязательно исключены посредством модификации методики измерения или замены приборов.
4. Наконец, нельзя забывать о возможности существования ошибок, о которых мы не подозреваем, но которые могут существенно искажать результаты измерений. Такие погрешности самые опасные, а исключить их можно только многократной *независимой* проверкой измерений, разными методами и в разных условиях.

В учебном практикуме учёт систематических погрешностей ограничивается, как правило, паспортными погрешностями приборов и теоретическими поправками к упрощенной модели исследуемого явления.

## 2. ЭЛЕМЕНТЫ ТЕОРИИ ОШИБОК

Результат любого измерения не определён однозначно и имеет случайную составляющую. Поэтому адекватным языком для описания погрешностей является язык вероятностей. Тот факт, что значение некоторой величины «случайно», не означает, что она может принимать совершенно произвольные значения. Ясно, что частоты, с которыми возникает те или иные значения, различны. Вероятностные законы, которым подчиняются случайные величины, называют *распределениями*.

### 2.1. Случайная величина

Назовём *вероятностью*  $P_A$  некоторого события  $A$  (например, результата измерения) *долю случаев*, в которых реализуется данный результат

в пределе большого числа измерений  $n$ :

$$P_A = \lim_{n \rightarrow \infty} \frac{n_A}{n},$$

где  $n_A$  — количество наблюдений результата  $A$ .

*Случайной* будем называть величину, значение которой не может быть *достоверно* определено экспериментатором с абсолютной точностью или значение которой меняется в каждом последующем опыте неконтролируемым образом. Согласно такому определению, *любой* результат любых измерений является случайным, поскольку осведомленность экспериментатора всегда чем-то ограничена.

**Замечание.** Хотя понятия вероятности и случайной величины являются основополагающими, в литературе нет единства в их определении. Обсуждение формальных тонкостей или построение строгой теории лежит за пределами данного пособия. Наша цель — дать упрощенное, но наглядное представление о предмете. Заинтересованным читателям рекомендуем обратиться к специальной литературе: [5].

Рассмотрим случайную физическую величину  $x$ , которая при измерениях может принимать *непрерывный* набор значений. Пусть  $P_{[x_0, x_0 + \delta x]}$  — вероятность того, что результат окажется вблизи некоторой точки  $x_0$  в пределах интервала  $\delta x$ :  $x \in [x_0, x_0 + \delta x]$ . Устремим интервал  $\delta x$  к нулю. Нетрудно понять, что вероятность попасть в этот интервал также будет стремиться к нулю. Однако отношение  $w(x_0) = \frac{P_{[x_0, x_0 + \delta x]}}{\delta x}$  будет оставаться конечным. Функцию  $w(x)$  называют *плотностью распределения вероятности* или кратко *распределением* непрерывной случайной величины  $x$ .

**Замечание.** В математической литературе распределением часто называют не функцию  $w(x)$ , а её интеграл  $W(x) = \int w(x) dx$ . Такую функцию в физике принято называть *интегральным* или *кумулятивным* распределением. В англоязычной литературе для этих функций принято использовать сокращения: *pdf* (*probability distribution function*) и *cdf* (*cumulative distribution function*) соответственно.

**Гистограммы.** Проиллюстрируем наглядно понятие плотности распределения. Результат большого числа измерений случайной величины удобно представить с помощью специального типа графика — *гистограммы*. Для этого область значений  $x$ , размещённую на оси абсцисс, разобьём на равные малые интервалы — «корзины» или «бины» (*англ.* bins) некоторого размера  $h$ . По оси ординат будет откладываться доля измерений  $w$ , результаты которых попадают в соответствующую корзину. А именно, пусть  $k$  — номер корзины;  $n_k$  — число измерений, попавших в диапазон  $x \in [kh, (k+1)h]$ . Тогда на графике изобразим «столбик» шириной  $h$  и вы-

сотой  $w_k = n_k/n$ . В результате получим картину, подобную изображённой на рис. 2.1.

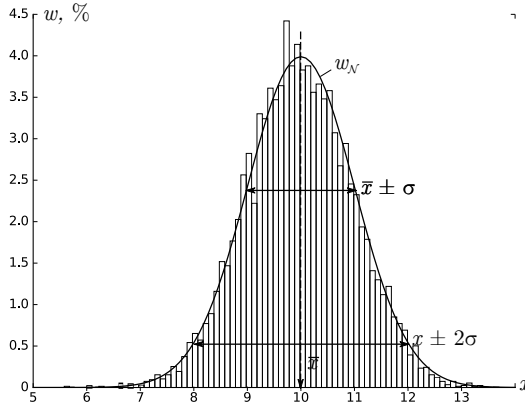


Рис. 2.1. Пример гистограммы для нормального распределения ( $\bar{x} = 10$ ,  $\sigma = 1,0$ ,  $h = 0,1$ ,  $n = 10^4$ )

Высоты построенных столбиков будут приблизительно соответствовать значению плотности распределения  $w(x)$  вблизи соответствующей точки  $x$ . Если устремить число измерений к бесконечности ( $n \rightarrow \infty$ ), а ширину корзины к нулю ( $h \rightarrow 0$ ), то огибающая гистограммы будет стремиться к некоторой непрерывной функции  $w(x)$ .

Самые высокие столбики гистограммы будут группироваться вблизи максимума функции  $w(x)$  — это *наиболее вероятное* значение случайной величины. Если отклонения в положительную и отрицательную стороны равновероятны, то гистограмма будет симметрична — в таком случае среднее значение  $\langle x \rangle$  также будет лежать вблизи этого максимума. Ширина гистограммы будет характеризовать разброс значений случайной величины — по порядку величины она, как правило, близка к среднеквадратичному отклонению  $s_x$ .

**Свойства распределений.** Из определения функции  $w(x)$  следует, что вероятность получить в результате эксперимента величину  $x$  в диапазоне от  $a$  до  $b$  можно найти, вычислив интеграл:

$$P_{x \in [a,b]} = \int_a^b w(x) dx. \quad (2.1)$$

Согласно определению вероятности, сумма вероятностей для всех возможных случаев всегда равна единице. Поэтому интеграл распределения

$w(x)$  по всей области значений  $x$  (то есть суммарная площадь под графиком  $w(x)$ ) равен единице:

$$\int_{-\infty}^{+\infty} w(x) dx = 1.$$

Это соотношение называют *условием нормировки*.

**Среднее и дисперсия.** Вычислим среднее по построенной гистограмме. Если размер корзины  $h$  достаточно мал, все измерения в пределах одной корзины можно считать примерно одинаковыми. Тогда среднее арифметическое всех результатов можно вычислить как

$$\langle x \rangle \approx \frac{1}{n} \sum_i n_i x_i = \sum_i w_i x_i.$$

Переходя к пределу, получим следующее определение среднего значения случайной величины:

$$\bar{x} = \int x w dx, \tag{2.2}$$

где интегрирование ведётся по всей области значений  $x$ . В теории вероятностей  $\bar{x}$  также называют *математическим ожиданием* распределения. Величину

$$\sigma^2 = \overline{(x - \bar{x})^2} = \int (x - \bar{x})^2 w dx \tag{2.3}$$

называют *дисперсией* распределения. Значение  $\sigma$  есть среднеквадратичное отклонение в пределе  $n \rightarrow \infty$ . Оно имеет ту же размерность, что и сама величина  $x$  и характеризует разброс распределения. Именно эту величину, как правило, приводят как характеристику погрешности измерения  $x$ .

**Доверительный интервал.** Обозначим как  $P_{|\Delta x| < \delta}$  вероятность того, что отклонение от среднего  $\Delta x = x - \bar{x}$  составит величину, не превосходящую по модулю значение  $\delta$ :

$$P_{|\Delta x| < \delta} = \int_{\bar{x}-\delta}^{\bar{x}+\delta} w(x) dx. \tag{2.4}$$

Эту величину называют *доверительной вероятностью* для *доверительного интервала*  $|x - \bar{x}| \leq \delta$ .



## 2.2. Нормальное распределение

Одним из наиболее примечательных результатов теории вероятностей является так называемая *центральная предельная теорема*. Она утверждает, что сумма большого количества независимых случайных слагаемых, каждое из которых вносит в эту сумму относительно малый вклад, подчиняется универсальному закону, не зависимо от того, каким вероятностным законам подчиняются её составляющие, — так называемому *нормальному распределению* (или *распределению Гаусса*).

Доказательство теоремы довольно громоздко и мы его не приводим (его можно найти в любом учебнике по теории вероятностей). Остановимся кратко на том, что такое нормальное распределение и его основных свойствах.

Плотность нормального распределения выражается следующей формулой:

$$w_{\mathcal{N}}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}. \quad (2.5)$$

Здесь  $\bar{x}$  и  $\sigma$  — параметры нормального распределения:  $\bar{x}$  равно среднему значению  $x$ , а  $\sigma$  — среднеквадратичному отклонению, вычисленным в пределе  $n \rightarrow \infty$ .

Как видно из рис. 2.1, распределение представляет собой симметричный «колокол», положение вершины которого соответствует  $\bar{x}$  (ввиду симметрии оно же совпадает с наиболее вероятным значением — максимумом функции  $w_{\mathcal{N}}(x)$ ).

При значительном отклонении  $x$  от среднего величина  $w_{\mathcal{N}}(x)$  очень быстро убывает. Это означает, что вероятность встретить отклонения, существенно большие, чем  $\sigma$ , оказывается *пренебрежимо мала*. Ширина «колокола» по порядку величины равна  $\sigma$  — она характеризует «разброс» экспериментальных данных относительно среднего значения.

**Замечание.** Точки  $x = \bar{x} \pm \sigma$  являются точками перегиба графика  $w(x)$  (в них вторая производная по  $x$  обращается в нуль,  $w'' = 0$ ), а их положение по высоте составляет  $w(\bar{x} \pm \sigma)/w(\bar{x}) = e^{-1/2} \approx 0,61$  от высоты вершины.

Универсальный характер центральной предельной теоремы позволяет широко применять на практике нормальное (гауссово) распределение для обработки результатов измерений, поскольку часто случайные погрешности складываются из множества случайных *независимых* факторов. Заметим, что на практике для *приближённой оценки* параметров нормального распределения случайной величины используются *выборочные* значения среднего и дисперсии:  $\bar{x} \approx \langle x \rangle$ ,  $s_x \approx \sigma_x$ .

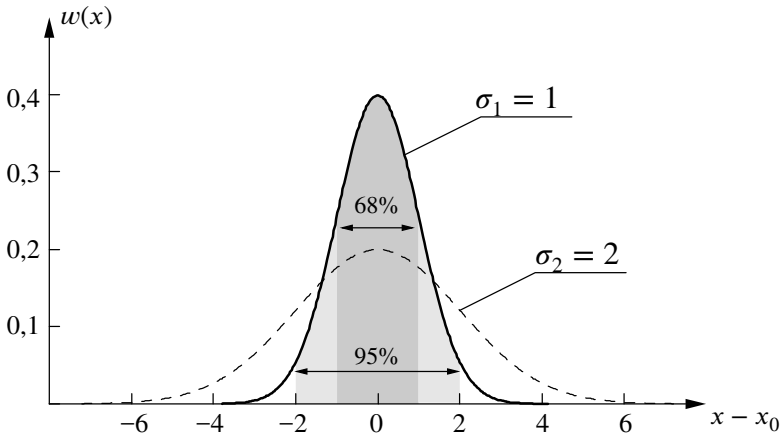


Рис. 2.2. Плотность нормального распределения

**Доверительные вероятности.** Вычислим некоторые доверительные вероятности (2.4) для нормально распределённых случайных величин.

**Замечание.** Значение интеграла вида  $\int e^{-x^2/2} dx$  (его называют *интегралом ошибок*) в элементарных функциях не выражается, но легко находится численно.

Вероятность того, что результат отдельного измерения  $x$  окажется в пределах  $\bar{x} \pm \sigma$  оказывается равна

$$P_{|\Delta x| < \sigma} = \int_{\bar{x} - \sigma}^{\bar{x} + \sigma} w_{\mathcal{A}} dx \approx 0,68.$$

Вероятность отклонения в пределах  $\bar{x} \pm 2\sigma$ :

$$P_{|\Delta x| < 2\sigma} \approx 0,95,$$

а в пределах  $\bar{x} \pm 3\sigma$ :

$$P_{|\Delta x| < 3\sigma} \approx 0,9973.$$

Иными словами, при большом числе измерений нормально распределённой величины можно ожидать, что лишь треть измерений выпадут за пределы интервала  $[\bar{x} - \sigma, \bar{x} + \sigma]$ . При этом около 5% измерений выпадут за пределы  $[\bar{x} - 2\sigma, \bar{x} + 2\sigma]$ , и лишь 0,27% окажутся за пределами  $[\bar{x} - 3\sigma, \bar{x} + 3\sigma]$ .

**Пример.** В сообщениях об открытии бозона Хиггса на Большом адронном коллайдере говорилось о том, что исследователи ждали подтверждение результатов с точностью «5 сигма». Используя нормальное распределение (2.5) нетрудно посчитать, что они использовали доверительную вероятность  $P \approx 1 - 5,7 \cdot 10^{-7} = 0,99999943$ . Такую точность можно назвать фантастической.

Полученные значения доверительных вероятностей используются при *стандартной записи результатов измерений*. В физических измерениях (в частности, в учебной лаборатории), как правило, используется  $P = 0,68$ , то есть, запись

$$x = \bar{x} \pm \delta x$$

означает, что измеренное значение лежит в диапазоне (доверительном интервале)  $x \in [\bar{x} - \delta x; \bar{x} + \delta x]$  с вероятностью 68%. Таким образом погрешность  $\pm \delta x$  считается равной одному среднеквадратичному отклонению:  $\delta x = \sigma$ . В *технических* измерениях чаще используется  $P = 0,95$ , то есть под абсолютной погрешностью имеется в виду удвоенное среднеквадратичное отклонение,  $\delta x = 2\sigma$ . Во избежание разночтений доверительную вероятность следует указывать отдельно.

**Замечание.** Хотя нормальный закон распределения встречается на практике довольно часто, стоит помнить, что он реализуется *далеко не всегда*. Полученные выше соотношения для вероятностей попадания значений в доверительные интервалы можно использовать в качестве простейшего признака нормальности распределения: в частности, если количество попадающих в интервал  $\pm \sigma$  результатов существенно отличается от  $2/3$  — это повод для более детального исследования закона распределения ошибок.

**Сравнение результатов измерений.** Теперь мы можем дать количественный критерий для сравнения двух измеренных величин или двух результатов измерения одной и той же величины.

Пусть  $x_1$  и  $x_2$  ( $x_1 \neq x_2$ ) измерены с погрешностями  $\sigma_1$  и  $\sigma_2$  соответственно. Ясно, что если различие результатов  $|x_2 - x_1|$  невелико, его можно объяснить просто случайными отклонениями. Если же теория предсказывает, что вероятность обнаружить такое отклонение слишком мала, различие результатов следует признать *значимым*. Предварительно необходимо договориться о соответствующем граничном значении вероятности. Универсального значения здесь быть не может, поэтому приходится полагаться на субъективный выбор исследователя. Часто в качестве «разумной» границы выбирают вероятность 5%, что, как видно из изложенного выше, для нормального распределения соответствует отклонению более, чем на  $2\sigma$ .

Допустим, одна из величин известна с существенно большей точностью:  $\sigma_2 \ll \sigma_1$  (например,  $x_1$  — результат, полученный студентом в лабо-

ратории,  $x_2$  — справочное значение). Поскольку  $\sigma_2$  мало,  $x_2$  можно принять за «истинное»:  $x_2 \approx \bar{x}$ . Предполагая, что погрешность измерения  $x_1$  подчиняется нормальному закону с дисперсией  $\sigma_1^2$ , можно утверждать, что различие считается значимым, если

$$|x_1 - x_2| > 2\sigma_1.$$

Пусть погрешности измерений сравнимы по порядку величины:  $\sigma_1 \sim \sigma_2$ . В теории вероятностей показывается, что линейная комбинация нормально распределённых величин также имеет нормальное распределение с дисперсией  $\sigma^2 = \sigma_1^2 + \sigma_2^2$  (см. также правила сложения погрешностей (2.7)). Тогда для проверки гипотезы о том, что  $x_1$  и  $x_2$  являются измерениями одной и той же величины, нужно вычислить, является ли значимым отклонение  $|x_1 - x_2|$  от нуля при  $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$ .

**Пример.** Два студента получили следующие значения для теплоты испарения некоторой жидкости:  $x_1 = 40,3 \pm 0,2$  кДж/моль и  $x_2 = 41,0 \pm 0,3$  кДж/моль, где погрешность соответствует одному стандартному отклонению. Можно ли утверждать, что они исследовали одну и ту же жидкость?

Имеем наблюдаемую разность  $|x_1 - x_2| = 0,7$  кДж/моль, среднеквадратичное отклонение для разности  $\sigma = \sqrt{0,2^2 + 0,3^2} = 0,36$  кДж/моль. Их отношение  $\frac{|x_2 - x_1|}{\sigma} \approx 2$ . Из свойств нормального распределения находим вероятность того, что измерялась одна и та же величина, а различия в ответах возникли из-за случайных ошибок:  $P \approx 5\%$ . Ответ на вопрос, «достаточно» ли мала или велика эта вероятность, остаётся на усмотрение исследователя.

**Замечание.** Изложенные здесь соображения применимы, только если  $\bar{x}$  и его стандартное отклонение  $\sigma$  получены на основании достаточно большой выборки  $n \gg 1$  (или заданы точно). При небольшом числе измерений ( $n \lesssim 10$ ) выборочные средние  $\langle x \rangle$  и среднеквадратичное отклонение  $s_x$  сами имеют довольно большую ошибку, а их распределение будет описываться не нормальным законом, а так называемым  $t$ -распределением Стьюдента. В частности, в зависимости от значения  $n$  интервал  $\langle x \rangle \pm s_x$  будет соответствовать несколько меньшей доверительной вероятности, чем  $P = 0,68$ . Особенно резко различия проявляются при высоких уровнях доверительных вероятностей  $P \rightarrow 1$ .

### 2.3. Независимые величины

Величины  $x$  и  $y$  называют *независимыми* если результат измерения одной из них никак не влияет на результат измерения другой. Для таких величин вероятность того, что  $x$  окажется в некоторой области  $X$ , и одновременно  $y$  — в области  $Y$ , равна произведению соответствующих

вероятностей:

$$P_{x \in X, y \in Y} = P_{x \in X} \cdot P_{y \in Y}.$$

Обозначим отклонения величин от их средних как  $\Delta x = x - \bar{x}$  и  $\Delta y = y - \bar{y}$ . Средние значения этих отклонений равны, очевидно, нулю:  $\overline{\Delta x} = \bar{x} - \bar{x} = 0$ ,  $\overline{\Delta y} = 0$ . Из независимости величин  $x$  и  $y$  следует, что среднее значение от произведения  $\overline{\Delta x \cdot \Delta y}$  равно произведению средних  $\overline{\Delta x} \cdot \overline{\Delta y}$  и, следовательно, равно нулю:

$$\overline{\Delta x \cdot \Delta y} = \overline{\Delta x} \cdot \overline{\Delta y} = 0. \quad (2.6)$$

### 2.3.1. Дисперсия суммы независимых величин

Пусть измеряемая величина  $z = x + y$  складывается из двух *независимых* случайных слагаемых  $x$  и  $y$ , для которых известны средние  $\bar{x}$  и  $\bar{y}$ , и их среднеквадратичные погрешности  $\sigma_x$  и  $\sigma_y$ . Непосредственно из определения (1.1) следует, что среднее суммы равно сумме средних:

$$\bar{z} = \bar{x} + \bar{y}.$$

Найдём дисперсию  $\sigma_z^2$ . В силу независимости имеем

$$\overline{\Delta z^2} = \overline{\Delta x^2} + \overline{\Delta y^2} + 2\overline{\Delta x \cdot \Delta y} \approx \overline{\Delta x^2} + \overline{\Delta y^2},$$

то есть:

$$\sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2. \quad (2.7)$$

Таким образом, при сложении *независимых* величин их погрешности складываются среднеквадратичным образом.

Подчеркнём, что для справедливости соотношения (2.7) величины  $x$  и  $y$  не обязаны быть нормально распределёнными — достаточно существования конечных значений их дисперсий. Однако можно показать, что если  $x$  и  $y$  распределены нормально, *нормальным будет и распределение их суммы*.

**Замечание.** Требование независимости слагаемых является принципиальным. Например, положим  $y = x$ . Тогда  $z = 2x$ . Здесь  $y$  и  $x$ , очевидно, зависят друг от друга. Используя (2.7), находим  $\sigma_{2x} = \sqrt{2}\sigma_x$ , что, конечно, неверно — непосредственно из определения следует, что  $\sigma_{2x} = 2\sigma_x$ .

Отдельно стоит обсудить математическую структуру формулы (2.7). Если одна из погрешностей много больше другой, например,  $\sigma_x \gg \sigma_y$ , то меньшей погрешностью можно пренебречь,  $\sigma_{x+y} \approx \sigma_x$ . С другой стороны, если два источника погрешностей имеют один порядок  $\sigma_x \sim \sigma_y$ , то и  $\sigma_{x+y} \sim \sigma_x \sim \sigma_y$ .

Эти обстоятельства важны при планировании эксперимента: как правило, величина, измеренная наименее точно, вносит наибольший вклад в погрешность конечного результата. При этом, пока не устранены наиболее существенные ошибок, бессмысленно гнаться за повышением точности измерения остальных величин.

**Пример.** Пусть  $\sigma_y = \sigma_x/3$ , тогда  $\sigma_z = \sigma_x \sqrt{1 + \frac{1}{9}} \approx 1,05\sigma_x$ , то есть при различии двух погрешностей более, чем в 3 раза, поправка к погрешности составляет менее 5%, и уже нет особого смысла в учёте меньшей погрешности:  $\sigma_z \approx \sigma_x$ . Это утверждение касается сложения любых независимых источников погрешностей в эксперименте.

## 2.4. Погрешность среднего

Выборочное среднее арифметическое значение  $\langle x \rangle$ , найденное по результатам  $n$  измерений, само является случайной величиной. Действительно, если поставить серию одинаковых опытов по  $n$  измерений, то в каждом опыте получится своё среднее значение, отличающееся от предельного среднего  $\bar{x}$ .

Вычислим среднеквадратичную погрешность среднего арифметического  $\sigma_{\langle x \rangle}$ . Рассмотрим вспомогательную сумму  $n$  слагаемых

$$Z = x_1 + x_2 + \dots + x_n.$$

Если  $\{x_i\}$  есть набор *независимых* измерений *одной и той же* физической величины, то мы можем, применяя результат (2.7) предыдущего параграфа, записать

$$\sigma_Z = \sqrt{\sigma_{x_1}^2 + \sigma_{x_2}^2 + \dots + \sigma_{x_n}^2} = \sqrt{n}\sigma_x,$$

поскольку под корнем находится  $n$  одинаковых слагаемых. Отсюда с учётом  $\langle x \rangle = Z/n$  получаем

$$\sigma_{\langle x \rangle} = \frac{\sigma_x}{\sqrt{n}}. \quad (2.8)$$

Таким образом, *погрешность среднего значения  $x$  по результатам  $n$  независимых измерений оказывается в  $\sqrt{n}$  раз меньше погрешности отдельного измерения.* Это один из важнейших результатов, позволяющий уменьшать случайные погрешности эксперимента за счёт многократного повторения измерений.

Подчеркнём отличия между  $\sigma_x$  и  $\sigma_{\langle x \rangle}$ :

величина  $\sigma_x$  — *погрешность отдельного измерения* — является характеристикой разброса значений в совокупности измерений  $\{x_i\}$ ,  $i = 1..n$ .

При нормальном законе распределения примерно 68% измерений попадают в интервал  $\langle x \rangle \pm \sigma_x$ ;

величина  $\sigma_{\langle x \rangle}$  — *погрешность среднего* — характеризует точность, с которой определено среднее значение измеряемой физической величины  $\langle x \rangle$  относительно предельного («истинного») среднего  $\bar{x}$ ; при этом с доверительной вероятностью  $P = 68\%$  искомая величина  $\bar{x}$  лежит в интервале  $\langle x \rangle - \sigma_{\langle x \rangle} < \bar{x} < \langle x \rangle + \sigma_{\langle x \rangle}$ .

## 2.5. Результирующая погрешность опыта

Пусть для некоторого результата измерения известна оценка его максимальной систематической погрешности  $\Delta_{\text{сист}}$  и случайная среднеквадратичная погрешность  $\sigma_{\text{случ}}$ . Какова «полная» погрешность измерения?

Предположим для простоты, что измеряемая величина *в принципе* может быть определена сколь угодно точно, так что можно говорить о некотором её «истинном» значении  $x_{\text{ист}}$  (иными словами, погрешность результата связана в основном именно с процессом измерения). Назовём *полной погрешностью* измерения среднеквадратичное значения отклонения от результата измерения от «истинного»:

$$\sigma_{\text{полн}}^2 = \left\langle (x - x_{\text{ист}})^2 \right\rangle.$$

Отклонение  $x - x_{\text{ист}}$  можно представить как сумму случайного отклонения от среднего  $\delta x_{\text{случ}} = x - \bar{x}$  и постоянной (но, вообще говоря, неизвестной) систематической составляющей  $\delta x_{\text{сист}} = \bar{x} - x_{\text{ист}} = \text{const}$ :

$$x - x_{\text{ист}} = \delta x_{\text{сист}} + \delta x_{\text{случ}}.$$

Причём случайную составляющую можно считать независимой от систематической. В таком случае из (2.7) находим:

$$\sigma_{\text{полн}}^2 = \langle \delta x_{\text{сист}}^2 \rangle + \langle \delta x_{\text{случ}}^2 \rangle \leq \Delta_{\text{сист}}^2 + \sigma_{\text{случ}}^2. \quad (2.9)$$

Таким образом, для получения *максимального* значения полной погрешности некоторого измерения нужно квадратично сложить максимальную систематическую и случайную погрешности.

Если измерения проводятся многократно, то согласно (2.8) случайная составляющая погрешности может быть уменьшена, а систематическая составляющая при этом остаётся неизменной:

$$\sigma_{\text{полн}}^2 \leq \Delta_{\text{сист}}^2 + \frac{\sigma_x^2}{n}.$$

Отсюда следует важное практическое правило (см. также обсуждение в п. 2.3.1): если случайная погрешность измерений в 2–3 раза меньше

предполагаемой систематической, то *нет смысла проводить многократные измерения* в попытке уменьшить погрешность всего эксперимента. В такой ситуации измерения достаточно повторить 2–3 раза — чтобы убедиться в повторяемости результата, исключить промахи и проверить, что случайная ошибка действительно мала. В противном случае повторение измерений может иметь смысл до тех пор, пока погрешность среднего  $\sigma_{\langle x \rangle} = \frac{\sigma_x}{\sqrt{n}}$  не станет меньше систематической.

**Замечание.** Поскольку конкретная величина систематической погрешности, как правило, не известна, её можно в некотором смысле рассматривать наравне со случайной — предположить, что её величина была определена по некоторому случайному закону перед началом измерений (например, при изготовлении линейки на заводе произошло некоторое случайное искажение шкалы). При такой трактовке формулу (2.9) можно рассматривать просто как частный случай формулы сложения погрешностей независимых величин (2.7).

Подчеркнем, что вероятностный закон, которому подчиняется систематическая ошибка, зачастую неизвестен. Поэтому неизвестно и распределение итогового результата. Из этого, в частности, следует, что мы не можем приписать интервалу  $x \pm \Delta_{\text{сист}}$  какую-либо определённую доверительную вероятность — она равна 0,68 только если систематическая ошибка имеет нормальное распределение. Можно, конечно, *предположить*, — и так часто делают — что, к примеру, ошибки при изготовлении линеек на заводе имеют гауссов характер. Также часто предполагают, что систематическая ошибка имеет *равномерное* распределение (то есть «истинное» значение может с равной вероятностью принять любое значение в пределах интервала  $\pm \Delta_{\text{сист}}$ ). Строго говоря, для этих предположений нет достаточных оснований.

**Пример.** В результате измерения диаметра проволоки микрометрическим винтом, имеющим цену деления  $h = 0,01$  мм, получен следующий набор из  $n = 8$  значений:

$d$ , мм	0,39	0,38	0,39	0,37	0,40	0,39	0,38	0,39
----------	------	------	------	------	------	------	------	------

Вычисляем среднее значение:  $\langle d \rangle \approx 386,3$  мкм. Среднеквадратичное отклонение:  $\sigma_d \approx 9,2$  мкм. Случайная погрешность среднего согласно (2.8):  $\sigma_{\langle d \rangle} = \frac{\sigma_d}{\sqrt{8}} \approx 3,2$  мкм. Все результаты лежат в пределах  $\pm 2\sigma_d$ , поэтому нет причин сомневаться в нормальности распределения. Максимальную погрешность микрометра оценим как половину цены деления,  $\Delta = \frac{h}{2} = 5$  мкм. Результирующая полная погрешность  $\sigma \leq \sqrt{\Delta^2 + \frac{\sigma_d^2}{8}} \approx 6,0$  мкм. Видно, что  $\sigma_{\text{случ}} \approx \Delta_{\text{сист}}$  и проводить дополнительные измерения особого смысла нет. Окончательно результат измерений может быть представлен в виде (см. та-



кже *правила округления* результатов измерений в п. 4.3.2)

$$d = 386 \pm 6 \text{ мкм}, \quad \varepsilon_d = 1,5\%.$$

Заметим, что поскольку случайная погрешность и погрешность прибора здесь имеют один порядок величины, наблюдаемый случайный разброс данных может быть связан как с неоднородностью сечения проволоки, так и с дефектами микрометра (например, с неровностями зажимов, люфтом винта, сухим трением, деформацией проволоки под действием микрометра и т. п.). Для ответа на вопрос, что именно вызвало разброс, требуются дополнительные исследования, желательно с использованием более точных приборов.

**Пример.** Измерение скорости полёта пули было осуществлено с погрешностью  $\delta v = \pm 1$  м/с. Результаты измерений для  $n = 6$  выстрелов представлены в таблице:

$v, \text{ м/с}$	146	170	160	181	147	168
------------------	-----	-----	-----	-----	-----	-----

Усреднённый результат  $\langle v \rangle = 162,0$  м/с, среднеквадратичное отклонение  $\sigma_v = 13,8$  м/с, случайная ошибка для средней скорости  $\sigma_{\bar{v}} = \sigma_v / \sqrt{6} = 5,6$  м/с. Поскольку разброс экспериментальных данных существенно превышает погрешность каждого измерения,  $\sigma_v \gg \delta v$ , он почти наверняка связан с реальным различием скоростей пули в разных выстрелах, а не с ошибками измерений. В качестве результата эксперимента представляют интерес как среднее значение скоростей  $\langle v \rangle = 162 \pm 6$  м/с ( $\varepsilon \approx 4\%$ ), так и значение  $\sigma_v \approx 14$  м/с, характеризующее разброс значений скоростей от выстрела к выстрелу. Малая инструментальная погрешность в принципе позволяет более точно измерить среднее и дисперсию, и исследовать закон распределения выстрелов по скоростям более детально — для этого требуется набрать большую статистику по выстрелам.

**Пример.** Измерение скорости полёта пули было осуществлено с погрешностью  $\delta v = 10$  м/с. Результаты измерений для  $n = 6$  выстрелов представлены в таблице:

$v, \text{ м/с}$	150	170	160	180	150	170
------------------	-----	-----	-----	-----	-----	-----

Усреднённый результат  $\langle v \rangle = 163,3$  м/с,  $\sigma_v = 12,1$  м/с,  $\sigma_{\langle v \rangle} = 5$  м/с,  $\sigma_{\text{полн}} \approx 11,2$  м/с. Инструментальная погрешность каждого измерения превышает разброс данных, поэтому в этом опыте затруднительно сделать вывод о различии скоростей от выстрела к выстрелу. Результат измерений скорости пули:  $\langle v \rangle = 163 \pm 11$  м/с,  $\varepsilon \approx 7\%$ . Проводить дополнительные выстрелы при такой большой инструментальной погрешности особого смысла нет — лучше поработать над точностью приборов и методикой измерений.

## 2.6. Обработка косвенных измерений

*Косвенными* называют измерения, полученные в результате расчётов, использующих результаты *прямых* (то есть «непосредственных») изме-

рений физических величин. Сформулируем основные правила пересчёта погрешностей при косвенных измерениях.

### 2.6.1. Случай одной переменной

Пусть в эксперименте измеряется величина  $x$ , а её «наилучшее» (в некотором смысле) значение равно  $x^*$  и оно известно с погрешностью  $\sigma_x$ . После чего с помощью известной функции вычисляется величина  $y = f(x)$ .

В качестве «наилучшего» приближения для  $y$  используем значение функции при «наилучшем»  $x$ :

$$y^* = f(x^*).$$

Найдём величину погрешности  $\sigma_y$ . Обозначая отклонение измеряемой величины как  $\Delta x = x - x^*$ , и пользуясь определением производной, при условии, что функция  $y(x)$  — гладкая вблизи  $x \approx x^*$ , запишем

$$\Delta y \equiv y(x) - y(x^*) \approx f' \cdot \Delta x,$$

где  $f' \equiv \frac{dy}{dx}$  — производная функции  $f(x)$ , взятая в точке  $x^*$ . Возведём полученное в квадрат, проведём усреднение ( $\sigma_y^2 = \langle \Delta y^2 \rangle$ ,  $\sigma_x^2 = \langle \Delta x^2 \rangle$ ), и затем снова извлечём корень. В результате получим

$$\sigma_y = \left| \frac{dy}{dx} \right| \sigma_x. \quad (2.10)$$

**Пример.** Для степенной функции  $y = Ax^n$  имеем  $\sigma_y = nAx^{n-1}\sigma_x$ , откуда

$$\frac{\sigma_y}{y} = n \frac{\sigma_x}{x}, \quad \text{или} \quad \varepsilon_y = n\varepsilon_x,$$

то есть относительная погрешность степенной функции возрастает пропорционально показателю степени  $n$ .

**Пример.** Для  $y = 1/x$  имеем  $\varepsilon_{1/x} = \varepsilon_x$  — при обращении величины сохраняется её относительная погрешность.

**Упражнение.** Найдите погрешность логарифма  $y = \ln x$ , если известны  $x$  и  $\sigma_x$ .

**Упражнение.** Найдите погрешность показательной функции  $y = a^x$ , если известны  $x$  и  $\sigma_x$ . Коэффициент  $a$  задан точно.

### 2.6.2. Случай многих переменных

Пусть величина  $u$  вычисляется по измеренным значениям нескольких различных *независимых* физических величин  $x, y, \dots$  на основе известного закона  $u = f(x, y, \dots)$ . В качестве наилучшего значения можно по-прежнему взять значение функции  $f$  при наилучших значениях измеряемых параметров:

$$u^* = f(x^*, y^*, \dots).$$

Для нахождения погрешности  $\sigma_u$  воспользуемся свойством, известным из математического анализа, — малые приращения гладких функции многих переменных складываются линейно, то есть справедлив *принцип суперпозиции* малых приращений:

$$\Delta u \approx f'_x \cdot \Delta x + f'_y \cdot \Delta y + \dots,$$

где символом  $f'_x \equiv \frac{\partial f}{\partial x}$  обозначена *частная производная* функции  $f$  по переменной  $x$  — то есть обычная производная  $f$  по  $x$ , взятая при условии, что все остальные аргументы (кроме  $x$ ) считаются постоянными параметрами. Тогда пользуясь формулой для нахождения дисперсии суммы независимых величин (2.7), получим соотношение, позволяющее вычислять погрешности косвенных измерений для произвольной функции  $u = f(x, y, \dots)$ :

$$\sigma_u^2 = f_x'^2 \sigma_x^2 + f_y'^2 \sigma_y^2 + \dots \quad (2.11)$$

Это и есть искомая общая формула пересчёта погрешностей при косвенных измерениях.

Отметим, что формулы (2.10) и (2.11) применимы только если относительные отклонения всех величин малы ( $\epsilon_x, \epsilon_y, \dots \ll 1$ ), а измерения проводятся вдали от особых точек функции  $f$  (производные  $f'_x, f'_y \dots$  не должны обращаться в бесконечность). Также подчеркнём, что все полученные здесь формулы справедливы только для *независимых* переменных  $x, y, \dots$

Остановимся на некоторых важных частных случаях формулы (2.11).

**Пример.** Для суммы (или разности)  $u = \sum_{i=1}^n a_i x_i$  имеем

$$\sigma_u^2 = \sum_{i=1}^n a_i^2 \sigma_{x_i}^2. \quad (2.12)$$

**Пример.** Найдём погрешность степенной функции:  $u = x^\alpha \cdot y^\beta \cdot \dots$ . Тогда нетрудно получить, что

$$\frac{\sigma_u^2}{u^2} = \alpha^2 \frac{\sigma_x^2}{x^2} + \beta^2 \frac{\sigma_y^2}{y^2} + \dots$$

или через относительные погрешности

$$\varepsilon_u^2 = \alpha^2 \varepsilon_x^2 + \beta^2 \varepsilon_y^2 + \dots \quad (2.13)$$

**Пример.** Вычислим погрешность произведения и частного:  $u = xy$  или  $u = x/y$ . Тогда в обоих случаях имеем

$$\varepsilon_u^2 = \varepsilon_x^2 + \varepsilon_y^2, \quad (2.14)$$

то есть при умножении или делении относительные погрешности складываются квадратично.

**Пример.** Рассмотрим несколько более сложный случай: нахождение угла по его тангенсу

$$u = \arctg \frac{y}{x}.$$

В таком случае, пользуясь тем, что  $(\arctg z)' = \frac{1}{1+z^2}$ , где  $z = y/x$ , и используя производную сложной функции, находим  $u'_x = u'_z z'_x = -\frac{y}{x^2+y^2}$ ,  $u'_y = u'_z z'_y = \frac{x}{x^2+y^2}$ , и наконец

$$\sigma_u^2 = \frac{y^2 \sigma_x^2 + x^2 \sigma_y^2}{(x^2 + y^2)^2}.$$

**Упражнение.** Найти погрешность вычисления гипотенузы  $z = \sqrt{x^2 + y^2}$  прямоугольного треугольника по измеренным катетам  $x$  и  $y$ .

По итогам данного раздела можно дать следующие практические рекомендации.

- Как правило, нет смысла увеличивать точность измерения какой-то одной величины, если другие величины, используемые в расчётах, остаются измеренными относительно грубо — всё равно итоговая погрешность скорее всего будет определяться самым неточным измерением. Поэтому все измерения имеет смысл проводить *примерно с одной и той же относительной погрешностью*.
- При этом, как следует из (2.13), особое внимание следует уделять измерению величин, возводимых при расчётах в степени с большими показателями. А при сложных функциональных зависимостях

имеет смысл детально проанализировать структуру формулы (2.11): если вклад от некоторой величины в общую погрешность мал, нет смысла гнаться за высокой точностью её измерения, и наоборот, точность некоторых измерений может оказаться критически важной.

- Следует избегать измерения малых величин как разности двух близких значений (например, толщины стенки цилиндра как разности внутреннего и внешнего радиусов): если  $u = x - y$ , то абсолютная погрешность  $\sigma_u = \sqrt{\sigma_x^2 + \sigma_y^2}$  меняется мало, однако относительная погрешность  $\varepsilon_u = \frac{\sigma_u}{x-y}$  может оказаться неприемлемо большой, если  $x \approx y$ .

### 3. ОЦЕНКА ПАРАМЕТРОВ

Цель любого физического эксперимента — проверить, выполняется ли некоторая теоретическая закономерность (*модель*), а также получить или уточнить её параметры. Поскольку набор экспериментальных данных неизбежно ограничен, а каждое отдельное измерение имеет погрешность, можно говорить лишь об *оценке* этих параметров. В большинстве случаев измеряется не одна величина, а некоторая функциональная зависимость величин друг от друга. В таком случае возникает необходимость построить оценку параметров этой зависимости.

**Пример.** Рассмотрим процедуру измерения сопротивления некоторого резистора. Простейшая теоретическая модель для резистора — закон Ома  $U = RI$ , где сопротивление  $R$  — единственный параметр модели. Часто при измерениях возможно возникновение систематической ошибки — смещение нуля напряжения или тока. Тогда для получения более корректной оценки сопротивления стоит использовать модель с двумя параметрами:  $U = RI + U_0$ .

Для построения оценки нужны следующие компоненты

- *данные* — результаты измерений  $\{x_i, y_i\}$  и их погрешности  $\{\sigma_i\}$  (экспериментальная погрешность является неотъемлемой частью набора данных!);
- *модель*  $y = f(x|\theta_1, \theta_2, \dots)$  — параметрическое описание исследуемой зависимости ( $\theta$  — набор параметров модели, например, коэффициенты  $\{k, b\}$  прямой  $f(x) = kx + b$ );
- процедура построения оценки параметров по измеренным данным («оценщик»):

$$\theta \approx \hat{\theta}(\{x_i, y_i, \sigma_i\}).$$

Рассмотрим самые распространенные способы построения оценки.

### 3.1. Метод минимума хи-квадрат

Обозначим отклонения результатов некоторой серии измерений от теоретической модели  $y = f(x|\theta)$  как

$$\Delta y_i = y_i - f(x_i|\theta), \quad i = 1 \dots n,$$

где  $\theta$  — некоторый параметр (или набор параметров), для которого требуется построить наилучшую оценку. Нормируем  $\Delta y_i$  на стандартные отклонения  $\sigma_i$  и построим сумму

$$\chi^2 = \sum_i \left( \frac{\Delta y_i}{\sigma_i} \right)^2, \quad (3.1)$$

которую принято называть суммой *хи-квадрат*.

Метод *минимума хи-квадрат* (*метод Пирсона*) заключается в подборе такого  $\theta$ , при котором сумма квадратов отклонений от теоретической модели, нормированных на ошибки измерений, достигает минимума:

$$\chi^2(\theta) \rightarrow \min.$$

**Замечание.** Подразумевается, что погрешность измерений  $\sigma_i$  указана только для вертикальной оси  $y$ . Поэтому, при использовании метода следует выбирать оси таким образом, чтобы относительная ошибка по оси абсцисс была значительно меньше, чем по оси ординат.

Данный метод вполне соответствует нашему интуитивному представлению о том, как теоретическая зависимость должна проходить через экспериментальные точки. Ясно, что чем ближе данные к модельной кривой, тем меньше будет сумма  $\chi^2$ . При этом, чем больше погрешность точки, тем в большей степени дозволено результатам измерений отклоняться от модели. Метода минимума  $\chi^2$  является частным случаем более общего *метода максимума правдоподобия* (см. ниже), реализующийся при *нормальном (гауссовом)* распределении ошибок.

Можно показать (см. [5]), что оценка по методу хи-квадрат является состоятельной, несмещенной и, если данные распределены нормально, имеет максимальную эффективность.

**Замечание.** Простые аналитические выражения для оценки методом хи-квадрат существуют (см. п. 3.6.1, 3.6.4) только в случае линейной зависимости  $f(x) = kx + b$  (впрочем, нелинейную зависимость часто можно заменой переменных свести к линейной). В общем случае задача поиска минимума

$\chi^2(\theta)$  решается численно, а соответствующая процедура реализована в большинстве специализированных программных пакетов по обработке данных.

### 3.2. Метод максимального правдоподобия.

Рассмотрим кратко один из наиболее общих методов оценки параметров зависимостей — метод максимума правдоподобия.

Сделаем два ключевых предположения:

- зависимость между измеряемыми величинами действительно может быть описана функцией  $y = f(x|\theta)$  при некотором  $\theta$ ;
- все отклонения  $\Delta y_i$  результатов измерений от теоретической модели являются *независимыми* и имеют *случайный* (не систематический!) характер.

Пусть  $P(\Delta y_i)$  — вероятность обнаружить отклонение  $\Delta y_i$  при фиксированных  $\{x_i\}$ , погрешностях  $\{\sigma_i\}$  и параметрах модели  $\theta$ . Построим функцию, равную вероятности обнаружить весь набор отклонений  $\{\Delta y_1, \dots, \Delta y_n\}$ . Ввиду независимости измерений она равна произведению вероятностей:

$$L = \prod_{i=1}^n P(\Delta y_i). \quad (3.2)$$

Функцию  $L$  называют *функцией правдоподобия*.

Метод максимума правдоподобия заключается в поиске такого  $\theta$ , при котором наблюдаемое отклонение от модели будет иметь *наибольшую вероятность*, то есть

$$L(\theta) \rightarrow \max.$$

**Замечание.** Поскольку с суммой работать удобнее, чем с произведениями, чаще используют не саму функцию  $L$ , а её логарифм:

$$\ln L = \sum_i \ln P(\Delta y_i).$$

Пусть теперь ошибки измерений имеют *нормальное* распределение (напомним, что согласно центральной предельной теореме нормальное распределение применимо, если отклонения возникают из-за большого числа независимых факторов, что на практике реализуется довольно часто). Согласно (2.5), вероятность обнаружить в  $i$ -м измерении отклонение  $\Delta y_i$  пропорциональна величине

$$P(\Delta y_i) \propto e^{-\frac{\Delta y_i^2}{2\sigma_i^2}},$$

где  $\sigma_i$  — стандартная ошибка измерения величины  $y_i$ . Тогда логарифм функции правдоподобия (3.2) будет равен (с точностью до константы)

$$\ln L = - \sum_i \frac{\Delta y_i^2}{2\sigma_i^2} = -\frac{1}{2}\chi^2.$$

Таким образом, максимум правдоподобия действительно будет соответствовать минимуму  $\chi^2$ .

### 3.3. Метод наименьших квадратов (МНК).

Рассмотрим случай, когда все погрешности измерений одинаковы,  $\sigma_i = \text{const}$ . Тогда множитель  $1/\sigma^2$  в сумме  $\chi^2$  выносится за скобки, и оценка параметра сводится к нахождению минимума суммы квадратов отклонений:

$$S(\theta) = \sum_{i=1}^n (y_i - f(x_i|\theta))^2 \rightarrow \min. \quad (3.3)$$

Оценка по *методу наименьших квадратов* (МНК) удобна в том случае, когда не известны погрешности отдельных измерений. Однако тот факт, что метод МНК игнорирует информацию о погрешностях, является и его основным недостатком. В частности, это не позволяет определить точность оценки (например, погрешности коэффициентов прямой  $\sigma_k$  и  $\sigma_b$ ) без привлечения дополнительных предположений (см. п. 3.6.2 и 3.6.3).

### 3.4. Проверка качества аппроксимации

Значение суммы  $\chi^2$  позволяет оценить, насколько хорошо данные описываются предлагаемой моделью  $y = f(x|\theta)$ .

Предположим, что распределение ошибок при измерениях *нормальное*. Тогда можно ожидать, что большая часть отклонений данных от модели будет порядка одной среднеквадратичной ошибки:  $\Delta y_i \sim \sigma_i$ . Следовательно, сумма хи-квадрат (3.1) окажется по порядку величины равна числу входящих в неё слагаемых:  $\chi^2 \sim n$ .

**Замечание.** Точнее, если функция  $f(x|\theta_1, \dots, \theta_p)$  содержит  $p$  подгоночных параметров (например,  $p = 2$  для линейной зависимости  $f(x) = kx + b$ ), то при заданных  $\theta$  лишь  $n - p$  слагаемых в сумме хи-квадрат будут независимы. Иными словами, когда параметры  $\theta$  определены из условия минимума хи-квадрат, сумму  $\chi^2$  можно рассматривать как функцию  $n - p$  переменных. Величину  $n - p$  называют *числом степеней свободы* задачи.

В теории вероятностей доказывается (см. [4] или [5]), что ожидаемое среднее значение (математическое ожидание) суммы  $\chi^2$  в точности равно числу степеней свободы:

$$\overline{\chi^2} = n - p.$$



Таким образом, при хорошем соответствии модели и данных, величина  $\chi^2/(n-p)$  должна в среднем быть равна единице. Значения существенно большие (2 и выше) свидетельствуют либо о *плохом соответствии теории и результатов измерений*, либо о *заниженных погрешностях*. Значения меньше 0,5 как правило свидетельствуют о *завышенных погрешностях*.

**Замечание.** Чтобы дать строгий количественный критерий, с какой долей вероятности гипотезу  $y = f(x)$  можно считать подтверждённой или опровергнутой, нужно знать вероятностный закон, которому подчиняется функция  $\chi^2$ . Если ошибки измерений распределены нормально, величина хи-квадрат подчиняется одноимённому распределению (с  $n-p$  степенями свободы). В элементарных функциях *распределение хи-квадрат* не выражается, но может быть легко найдено численно: функция встроена во все основные статистические пакеты, либо может быть вычислена по таблицам.

### 3.5. Оценка погрешности параметров

Важным свойством метода хи-квадрат является «встроенная» возможность нахождения погрешности вычисленных параметров  $\sigma_\theta$ .

Пусть функция  $L(\theta)$  имеет максимум при  $\theta = \hat{\theta}$ , то есть  $\hat{\theta}$  — решение задачи о максимуме правдоподобия. Согласно центральной предельной теореме мы ожидаем, что функция правдоподобия будем близка к нормальному распределению:  $L(\theta) \propto \exp\left(-\frac{(\theta-\hat{\theta})^2}{2\sigma_\theta^2}\right)$ , где  $\sigma_\theta$  — искомая погрешность параметра. Тогда в окрестности  $\hat{\theta}$  функция  $\chi^2(\theta) = -2 \ln(L(\theta))$  имеет вид параболы:

$$\chi^2(\theta) = \frac{(\theta - \hat{\theta})^2}{\sigma_\theta^2} + \text{const.}$$

Легко убедиться, что:

$$\chi^2(\hat{\theta} \pm \sigma_\theta) - \chi^2(\hat{\theta}) = 1.$$

Иными словами, при отклонении параметра  $\theta$  на одну ошибку  $\sigma_\theta$  от значения  $\hat{\theta}$ , минимизирующего  $\chi^2$ , функция  $\chi^2(\theta)$  изменится на единицу. Таким образом для нахождения *интервальной* оценки для искомого параметра достаточно графическим или численным образом решить уравнение

$$\Delta\chi^2(\theta) = 1. \tag{3.4}$$

Вероятностное содержание этого интервала будет равно 68% (его еще называют  $1-\sigma$  интервалом). Отклонение  $\chi^2$  на 2 будет соответствовать уже 95% доверительному интервалу.

**Замечание.** Приведенное решение просто использовать только в случае одного параметра. Впрочем, все приведенные рассуждения верны и в многопараметрическом случае. Просто решением уравнения 3.4 будет не отрезок, а некоторая многомерная фигура (эллипс в двумерном случае и гипер-эллипс при больших размерностях пространства параметров). Вероятностное содержание области, ограниченной такой фигурой будет уже не равно 68%, но может быть вычислено по соответствующим таблицам. Подробнее о многомерном случае в разделе 5.5.

### 3.6. Методы построения наилучшей прямой

Применим перечисленные выше методы к задаче о построении наилучшей прямой  $y = kx + b$  по экспериментальным точкам  $\{x_i, y_i\}$ . Линейность функции позволяет записать решение в относительно простом аналитическом виде.

Обозначим расстояние от  $i$ -й экспериментальной точки до искомой прямой, измеренное по вертикали, как

$$\Delta y_i = y_i - (kx_i + b),$$

и найдём такие параметры  $\{k, b\}$ , чтобы «совокупное» отклонение результатов от линейной зависимости было в некотором смысле минимально.

#### 3.6.1. Метод наименьших квадратов

Пусть сумма квадратов расстояний от точек до прямой минимальна:

$$S(k, b) = \sum_{i=1}^n (y_i - (kx_i + b))^2 \rightarrow \min. \quad (3.5)$$

Данный метод построения наилучшей прямой называют *методом наименьших квадратов* (МНК).

Рассмотрим сперва более простой частный случай, когда искомая прямая заведомо через «ноль», то есть  $b = 0$  и  $y = kx$ . Необходимое условие минимума функции  $S(k)$ , как известно, есть равенство нулю её производной. Дифференцируя сумму (3.5) по  $k$ , считая все величины  $\{x_i, y_i\}$  константами, найдём

$$\frac{dS}{dk} = - \sum_{i=1}^n 2x_i (y_i - kx_i) = 0.$$

Решая относительно  $k$ , находим

$$k = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.$$

Поделив числитель и знаменатель на  $n$ , этот результат можно записать более компактно:

$$k = \frac{\langle xy \rangle}{\langle x^2 \rangle}. \quad (3.6)$$

Напомним, что угловые скобки означают усреднение по всем экспериментальным точкам:

$$\langle \dots \rangle \equiv \frac{1}{n} \sum_{i=1}^n (\dots)_i$$

В общем случае при  $b \neq 0$  функция  $S(k, b)$  должна иметь минимум как по  $k$ , так и по  $b$ . Поэтому имеем систему из двух уравнений  $\partial S / \partial k = 0$ ,  $\partial S / \partial b = 0$ , решая которую, можно получить (получите самостоятельно):

$$k = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2}, \quad b = \langle y \rangle - k \langle x \rangle. \quad (3.7)$$

Эти соотношения и есть решение задачи о построении наилучшей прямой методом наименьших квадратов.

**Замечание.** Совсем кратко формулу (3.7) можно записать, если ввести обозначение

$$D_{xy} \equiv \langle xy \rangle - \langle x \rangle \langle y \rangle = \langle (x - \langle x \rangle) \cdot (y - \langle y \rangle) \rangle. \quad (3.8)$$

В математической статистике величину  $D_{xy}$  называют *ковариацией*. При  $x \equiv y$  имеем дисперсию  $D_{xx} = \langle (x - \langle x \rangle)^2 \rangle$ . Тогда

$$k = \frac{D_{xy}}{D_{xx}}, \quad b = \langle y \rangle - k \langle x \rangle. \quad (3.9)$$

### 3.6.2. Погрешность МНК в линейной модели

Погрешности  $\sigma_k$  и  $\sigma_b$  коэффициентов, вычисленных по формуле (3.7) (или (3.6)), можно оценить в следующих предположениях. Пусть погрешность измерений величины  $x$  пренебрежимо мала:  $\sigma_x \approx 0$ , а погрешности по  $y$  *одинаковы* для всех экспериментальных точек  $\sigma_y = \text{const}$ , *независимы* и имеют *случайный* характер (систематическая погрешность отсутствует).

Пользуясь в этих предположениях формулами для погрешностей косвенных измерений (см. раздел (2.6)) можно получить следующие соотношения:

$$\sigma_k = \sqrt{\frac{1}{n-2} \left( \frac{D_{yy}}{D_{xx}} - k^2 \right)}, \quad (3.10)$$

$$\sigma_b = \sigma_k \sqrt{\langle x^2 \rangle}, \quad (3.11)$$

где использованы введённые выше сокращённые обозначения (3.8). Коэффициент  $n - 2$  отражает число независимых «степеней свободы»:  $n$  экспериментальных точек за вычетом двух условий связи (3.7).

В частном случае  $y = kx$ :

$$\sigma_k = \sqrt{\frac{1}{n-1} \left( \frac{\langle y^2 \rangle}{\langle x^2 \rangle} - k^2 \right)}. \quad (3.12)$$

### 3.6.3. Недостатки и условия применимости МНК

Формулы (3.7) (или (3.6)) позволяют провести прямую по *любому* набору экспериментальных данных, а полученные выше соотношения — вычислить соответствующую среднеквадратичную ошибку для её коэффициентов. Однако далеко не всегда результат будет иметь физический смысл. Перечислим ограничения применимости данного метода.

В первую очередь метод наименьших квадратов — статистический, и поэтому он предполагает использование достаточно большого количества экспериментальных точек (желательно  $n > 10$ ).

Поскольку метод предполагает наличие погрешностей только по  $y$ , оси следует выбирать так, чтобы погрешность  $\sigma_x$  откладываемой по оси абсцисс величины была минимальна.

Кроме того, метод предполагает, что все погрешности в опыте — случайны. Соответственно, формулы (3.10)–(3.12) применимы *только для оценки случайной составляющей* ошибки  $k$  или  $b$ . Если в опыте предполагаются достаточно большие систематические ошибки, они должны быть оценены *отдельно*. Отметим, что для оценки систематических ошибок не существует строгих математических методов, поэтому в таком случае проще и разумнее всего воспользоваться графическим методом.

Одна из основных проблем, связанных с определением погрешностей методом наименьших квадратов заключается в том, что он дает разумные погрешности даже в том случае, когда данные вообще не соответствуют модели. Если погрешности измерений известны, предпочтительно использовать метод минимума  $\chi^2$ .

Наконец, стоит предостеречь от использования *любых* аналитических методов «вслепую», без построения графиков. В частности, МНК не способен выявить такие «аномалии», как отклонения от линейной зависимости, немонотонность, случайные всплески и т.п. Все эти случаи требуют особого рассмотрения и могут быть легко обнаружены визуально при построении графика.

### 3.6.4. Метод хи-квадрат построения прямой

Пусть справедливы те же предположения, что и для метода наименьших квадратов, но погрешности  $\sigma_i$  экспериментальных точек различны.

Метод минимума хи-квадрат сводится к минимизации суммы квадратов отклонений, где каждое слагаемое взято с *весом*  $w_i = 1/\sigma_i^2$ :

$$\chi^2(k,b) = \sum_{i=1}^n w_i (y_i - (kx_i + b))^2 \rightarrow \min.$$

Этот метод также называют *взвешенным* методом наименьших квадратов.

Определим *взвешенное среднее* от некоторого набора значений  $\{x_i\}$  как

$$\langle x \rangle' = \frac{1}{W} \sum_i w_i x_i,$$

где  $W = \sum_i w_i$  — нормировочная константа.

Повторяя процедуру, использованную при выводе (3.7), нетрудно получить (получите) совершенно аналогичные формулы для искомым коэффициентов:

$$k = \frac{\langle xy \rangle' - \langle x \rangle' \langle y \rangle'}{\langle x^2 \rangle' - \langle x \rangle'^2}, \quad b = \langle y \rangle' - k \langle x \rangle', \quad (3.13)$$

с тем отличием от (3.7), что под угловыми скобками  $\langle \dots \rangle'$  теперь надо понимать усреднение с весами  $w_i = 1/\sigma_i^2$ .

Записанные формулы позволяют вычислить коэффициенты прямой, *если* известны погрешности  $\sigma_{y_i}$ . Значения  $\sigma_{y_i}$  могут быть получены либо из некоторой теории, либо измерены непосредственно (многократным повторением измерений при каждом  $x_i$ ), либо оценены из каких-то дополнительных соображений (например, как инструментальная погрешность).

## 4. РЕКОМЕНДАЦИИ ПО ВЫПОЛНЕНИЮ И ПРЕДСТАВЛЕНИЮ РЕЗУЛЬТАТОВ РАБОТЫ

### 4.1. Проведение измерений

Ключевым элементом проведения лабораторной работы является ведение *лабораторного журнала*. Журнал является главным источником информации о проведенном эксперименте.

#### 4.1.1. Правила ведения лабораторного журнала

- Лабораторный журнал оформляется *строго от руки*. Для оформления лучше использовать большую тетрадь *формата А4* с несъемными листами. Это правило связано с тем, что никакой электронный журнал не обладает такой же информативностью и гибкостью в

оформлении. Рукописные журналы используются на всех крупных современных физических экспериментах.

- В журнале необходимо фиксировать всю информацию о проводимом эксперименте: название работы, дату и время проведения эксперимента, типы использованных приборов, схему установки, а также любые другие показатели, которые могут быть связаны с проведением работы и обработкой результатов.

**Замечание.** Недопустимым считается отсутствие какой-то информации в журнале поскольку она «есть в лабнике». Информация в описании может быть устаревшей или не соответствовать конкретной установке. Допускается использование элементов описания, нарисованных на компьютере, напечатанных и вклеенных в журнал.

- Лабораторный журнал должен содержать *максимально полную* информацию о процессе проведения эксперимента, а не только результаты измерений. Обязательно должны быть указаны все проводимые экспериментатором действия. По возможности должны присутствовать временные метки всех действий (например, чтобы потом можно было сверить журнал с журналами других студентов, работающих в это время, или с другой информацией).

**Пример.** Результаты измерений могут зависеть от окружающей температуры и влажности (особенно в работах по термодинамике). Поэтому если в момент измерений, кто-то открывает дверь или окно в лаборатории, может случиться синхронный скачок измеряемых значений на всех установках. Этот скачок может быть не замечен на стадии измерений и обнаружен только при обработке. Если хотя бы в одном журнале есть запись о том, что была открыта дверь, а во всех остальных есть временные метки, то можно при обработке учесть изменение условий.

- Не допускается исключение из журнала «неправильных» (или показавшихся неправильными) измерений. Если по какой-то причине сделано заключение о том, что измерение проведено в неправильных условиях, результаты должны быть сохранены, а в журнале сделана пометка о том, почему это измерение считается ненадежным. История знает много примеров, когда на первый взгляд «ошибочные» измерения приводили к открытиям.
- Рекомендуется дублировать в журнале показания приборов даже если они записываются автоматически электронным способом. Это позволяет избежать многих ошибок.

- При записи результатов измерений не допускается использование карандаша, корректора или черновиков.

#### 4.1.2. Подготовка к работе

Перед выполнением учебной лабораторной работы необходимо

- ознакомиться с описанием работы и теоретическим введением по соответствующей теме: получить таким образом представление об изучаемых явлениях, порядках измеряемых величин и связывающих их закономерностях, а также о методе измерения, используемых приборах и последовательности действий при проведении измерений;
- продумать предложенный в описании план действий, оценить необходимое количество измерений. Количество измерений студент должен оценивать самостоятельно исходя из а) требуемой точности измерений и б) планируемого времени выполнения работы;
- желательно заранее (в крайнем случае, на начальном этапе работы) представлять диапазон изменения измеряемых величин и выбрать для них соответствующие единицы измерения;
- предварительно оценить достижимую точность измерений, проанализировать возможные источники погрешностей и их влияние на погрешность конечного результата.

Для подготовки к выполнению работы рекомендуется наличие в журнале следующих элементов:

- название (не только номер!) и цели работы;
- схема установки и описание использованных приборов. Следует иметь в виду, что *реальная* схема конкретной установки может отличаться от той, что изображена в описании;
- основные теоретические положения и расчётные формулы для данной работы. Не следует переписывать (или перепечатывать) всё, что изложено в описании работы — нужно выделить ключевые моменты, необходимые для проведения работы и интерпретации результатов.
- план работы с оценкой количества измерений и времени, необходимого на выполнение каждого пункта. В процессе работы план может меняться, о чем должна быть сделана соответствующая пометка в журнале (с указанием причин).

### 4.1.3. Начало работы

В начале работы необходимо тщательно ознакомиться с экспериментальной установкой, проверить работоспособность приборов. Все сведения о приборах и условиях эксперимента необходимо зафиксировать в лабораторном журнале. Рекомендуется переписать полные наименования приборов — в этом случае недостающую информацию о них можно всегда найти в интернете.

**Замечание.** При сборке электрических схем источники питания подключаются к схеме в последнюю очередь. Регулировочные ручки напряжения или тока должны исходно находиться в *нулевом* положении.

Прежде чем приступить к основным измерениям, необходимо проверить работу установки. Первые измерения должны быть контрольными, чтобы убедиться, что все работает нормально, диапазон и точность измерений выбраны правильно. Если разброс повторных измерений не превышает инструментальную погрешность, то многократных измерений не требуется.

Замеченные неполадки в работе приборов и установок надо зафиксировать в журнале и сообщить об этом преподавателю.

### 4.1.4. Выбор количества измерений

Выбор количества измерений является сложной задачей, не имеющей единого алгоритма принятия решений. Тем не менее, каждый экспериментатор (в том числе, студент) должен самостоятельно определять, какое количество измерений является достаточным, базируясь на соображениях точности результатов, времени измерений и здравого смысла.

**Измерение фиксированной величины.** При измерении некоторой фиксированной величины количество необходимых измерений зависит от разброса результатов. Для первичной оценки этого разброса рекомендуется проделать измерения как минимум 3–4 раза (если позволяет время). Разброс полученных значений приблизительно соответствует статистической ошибке отдельного измерения. Если разброс существенно превышает точность измерительных приборов, то имеет смысл (опять же если позволяет время) провести более длительную серию (8–10) измерений, и после этого вычислить среднеквадратичное отклонение отдельного измерения от среднего.

Если остальные измерения в серии проводятся аналогичным образом, то разумно ожидать, что разброс остальных измерений будет таким же, и повторять длинную серию для всех измерений не нужно.



**Пример.** Допустим, требуется с помощью секундомера измерять периоды колебания маятника с точностью  $\varepsilon = 0,1\%$ . Предположим, что ошибка измерения связана только с временем реакции экспериментатора. Эта ошибка, очевидно, не зависит от длительности измерения и её можно измерить непосредственно: для этого можно 8–10 раз измерить время некоторого целого числа колебаний и по результатам вычислить среднеквадратичную погрешность времени реакции  $\sigma_t^{\text{реакц}}$  (как правило,  $\sigma_t^{\text{реакц}} \sim 0,2$  с).

По заданной абсолютной величине погрешности и требуемой точности  $\varepsilon$  находим необходимое полное время измерений:  $t = \sigma_t^{\text{реакц}}/\varepsilon \sim 200$  с. Тогда все последующие измерения можно не повторять многократно, а проводить 1–2 раза в течение рассчитанного времени  $t$ .

Эти рассуждения не учитывают возможное отставание или опережение часов при больших  $t$  — предполагается, что часы откалиброваны с достаточной точностью (их «уход» за время  $t$  не превышает времени реакции).

**Измерение зависимостей.** При измерениях функциональной зависимости в первую очередь следует позаботиться о том, насколько хорошо будут восстанавливаться параметры этой зависимости. Число параметров не может быть больше, чем число экспериментальных точек (нельзя строить прямую по одной точке!). Но даже в случае, если число точек равно числу параметров, эксперимент нельзя считать удовлетворительным, поскольку нет возможности проверить, является ли модель правильной и не было ли одно из измерений ошибочным. Универсального правила по выбору количества точек нет, но для определения параметров прямой рекомендуется иметь не менее 5–6 точек (а лучше 8–10).

В случае длительных измерений, следует заранее планировать время таким образом, чтобы точки максимально равномерно лежали во всем диапазоне измерений. Также важно понимать, что в случае измерения зависимостей, количество точек с разными параметрами важнее, чем точность отдельного измерения, поскольку при аппроксимации параметров модели, накопленная информация по разным точкам все равно будет просуммирована.

Важно отметить, что часто параметры установки «плывут» («дрейфуют») во время проведения эксперимента, поэтому рекомендуется при измерениях делать проходы в одну и в другую сторону по всему диапазону значений.

#### 4.1.5. Измерения

Результаты измерений и сопутствующих вычислений должны быть представлены в *таблицах*. Таблицы должны иметь подписи с кратким описанием их содержания и, возможно, с пояснениями по структуре расположения данных. Заглавные столбцы (или строки) должны быть подписаны, в них должны быть указаны буквенные обозначения величин

(введенные в тексте ранее) и их размерность.

При записи результатов измерений фиксируются *непосредственные показания прибора* — без какого либо пересчёта единиц измерения, округления и т.п. В частности, если прибор имеет шкалу, записывается *число делений* отклонения стрелки, и отдельно — цена деления (в отдельном столбце или перед таблицей). Пересчёт в физические единицы с учётом цены деления производится позже при обработке. Это позволяет минимизировать ошибки при снятии показаний.

Полезно строить предварительные графики (прямо в экспериментальном журнале) зависимостей измеряемых величин по мере получения результатов. При этом сразу выделяются области резких изменений, в которых измерения должны проводиться подробнее (больше точек), чем на участках плавного изменения. Если изучаемая закономерность, например линейная, выполняется только на некотором участке, то область измерений должна быть выбрана шире этого участка, чтобы можно было установить границы работы закономерности.

#### 4.1.6. Расчёты, анализ и представление результатов

Полученные первичные результаты в виде таблиц и графиков используются для расчёта конечных значений величин и их погрешностей либо для нахождения зависимости измеряемых величин между собой.

В общем случае лабораторный журнал и отчет — это *два разных документа* и могут быть оформлены по отдельности. Отчет не должен быть столь же подробен как журнал с точки зрения деталей проведения эксперимента. Также в отчете можно опустить прямые результаты измерений — достаточно дать ссылки на соответствующие страницы журнала.

Для измеряемых величин окончательные результаты должны быть представлены в виде среднего значения, погрешности и количества проведённых измерений.

Для окончательной оценки качества результатов необходимо сравнить их с данными, приводимыми в справочниках.

**Замечание.** Совпадение или несовпадение измеренного значения со справочным не может считаться критерием правильности проведения работы. Во-первых, значения действительно могут отличаться. Материалы, используемые в лабораторных работах, не всегда являются чистыми и соответствуют справочнику. Во-вторых, могут быть объективные причины, по которым результаты разошлись. Поиск и объяснения этих причин является более важным, чем точное совпадение значений!

## 4.2. Анализ инструментальных погрешностей

Перед выполнением любого эксперимента необходимо предварительно проанализировать возможные погрешности используемых приборов. Они

могут иметь как систематический, так и случайный характер. Можно говорить о единой оценке *инструментальной погрешности* прибора  $\sigma_{\text{инстр}}$ , которая учитывает обе составляющие.

**Погрешность шкалы.** При работе с прибором *со шкалой* (линейка, штангенциркуль, стрелочные приборы и т.д.) один из источников погрешности — необходимость выбора некоторого значения (интерполяции) между метками шкалы. Эта погрешность, которую как правило оценивают в *половину цены деления*, называется *погрешностью отсчёта* по шкале. Аналогичная погрешность есть и у приборов с цифровым дисплеем — это погрешность округления цифры последнего разряда. Данная погрешность может быть как случайной, так и систематической: в частности, если показания прибора стабильны (стрелка не дрожит и при повторных измерениях стрелка попадает в то же самое место шкалы), ошибка отсчёта будет систематической; если стрелка дрожит (или «плавает» последняя цифра разряда), ошибка будет случайной.

**Замечание.** Стоит по возможности избегать измерений в начале шкалы: если измеряемая величина лишь немногим превосходит цену деления (или единицу последнего разряда дисплея), относительная ошибка измерения резко возрастает.

**Паспортная погрешность.** Любой прибор имеет погрешность изготовления, калибровки, а также внутренние источники ошибок (например, шумы). Как правило, максимальные значения этих погрешностей определяются производителем и описаны в паспорте прибора. Погрешности могут зависеть от условий эксплуатации (температура, влажность и т.д.), что также должно отражаться в паспорте.

**Пример.** Согласно паспорту вольтметра В7–34, его относительная погрешность при работе на пределе измерений 1 В, оценивается по формуле

$$\varepsilon_x = \left[ 0,015 + 0,002 \left( \frac{1 \text{ В}}{U_x} - 1 \right) \right] \cdot [1 + 0,01 \cdot |t - 20|],$$

где  $U_x$  [В] — значение измеряемой величины,  $t$  [°С] — комнатная температура. Если измерения проводятся при температуре 24 °С и прибор показывает напряжение  $U_x = 500$  мВ, то относительная погрешность равна  $\varepsilon \approx 1,7\%$ , а абсолютная  $\delta U \approx \pm 8$  мВ.

Для стрелочных приборов традиционно используется понятие *класса точности*. Предельная инструментальная погрешность равна произведению класса точности (в процентах) на показание прибора при максимальном отклонении стрелки. В цифровых приборах погрешность, как правило, зависит от диапазона измерения, поэтому понятие класса точности для них не применяется.

**Пример.** Стрелочный вольтметр имеет диапазон измерения от 0 до 5 В и цену деления 10 мВ, а его класс точности равен 0,5. Следовательно, погрешность измерения, гарантируемая производителем, составляет  $5 \text{ В} \cdot 0,5\% = 25 \text{ мВ}$ . Хотя цена деления меньше, в качестве погрешности следует взять именно  $\pm 25 \text{ мВ}$ . Не стоит рассчитывать на хорошую точность при измерениях напряжения менее 1 В, поскольку относительная ошибка составит более 2,5%.

**Сложение погрешностей.** Наконец, при считывании показаний стрелка прибора или цифры на циферблате могут «дрожать» (*флуктуировать*) вблизи некоторого значения. Это может быть связано как с разного рода шумами и помехами внутри прибора, так и с колебаниями самой измеряемой величины. Если записывается некоторое среднее значение показаний, то амплитуда флуктуаций должна быть учтена как дополнительная случайная погрешность.

Не стоит также забывать, что в процессе эксперимента почти наверняка возникнут дополнительные погрешности, связанные с конкретной постановкой опыта и методикой измерений. Для нахождения результирующей погрешности измерения необходимо сложить все *независимые* источники ошибок среднеквадратичным образом:

$$\sigma_{\text{полн}} = \sqrt{\sigma_{\text{инстр}}^2 + \sigma_{\text{отсч}}^2 + \sigma_{\text{случ}}^2 + \dots}$$

**Замечание.** Отметим, что цену деления шкалы или разрядность дисплея *добросовестный* производитель выбирает таким образом, чтобы погрешность отсчёта и погрешность самого прибора были согласованы. В таком случае погрешность отсчёта по шкале отдельно учитывать не нужно — она уже учтена производителем при расчёте инструментальной погрешности.

### 4.3. Отчёт о работе

Лабораторная работа студента — миниатюрное научное исследование. Настоящие требования основаны на общепринятых стандартах научных публикаций, упрощенных для студентов младших курсов.

Отчёт о проделанной лабораторной работе должен представлять собой целостный документ, позволяющий читателю получить максимально полную информацию о проделанной работе и полученных результатах — без каких-либо дополнительных пояснений со стороны студента.

Материал в отчёте должен излагаться последовательно, а сам отчёт должен быть структурирован по разделам. Отчёт, как правило, содержит разделы: 1) аннотация, 2) теоретические сведения, 3) методика измерений, 4) используемое оборудование, 5) результаты измерений и обработка данных, 6) обсуждение результатов, 7) заключение. Структура и названия

разделов могут незначительно варьироваться в зависимости от конкретного содержания работы.

Записи лабораторного журнала прикрепляются к отчёту в качестве приложения. Допускается ведение лабораторного журнала и оформление отчётов в одной рабочей тетради (формата А4).

Размерность измеренных величин — как в таблицах, так и на графиках — должна быть подобрана так, чтобы данные были удобны для чтения и *не содержали избыточное количество нулей*.

Помимо таблиц и графиков в тексте отчёта также должны быть представлены промежуточные результаты обработки данных (с соответствующими погрешностями), указаны используемые методы обработки данных и приведены соответствующие формулы. Окончательные и наиболее важные промежуточные результаты должны быть записаны с указанием погрешности (как абсолютной, так и относительной) и округлены согласно принятым в физике правилам округления.

#### 4.3.1. Требования к содержанию разделов

**Аннотация:** краткое (1–2 абзаца) описание работы: её цели, используемые методы и приборы, ожидаемые результаты.

**Теоретические сведения:** краткий обзор основных понятий и теоретических законов, используемых или проверяемых в работе; упрощения и предположения, используемые при анализе и интерпретации результатов эксперимента; основные расчётные формулы.

**Методика измерений:** схема и описание экспериментальной установки; краткое описание основных методик проведения эксперимента, получения и обработки экспериментальных данных.

**Используемое оборудование:** перечень измерительных приборов, используемых в работе; инструментальные погрешности приборов и предварительный анализ их влияния на результаты опыта.

**Результаты измерений и обработка данных:** результаты проведенных измерений в форме таблиц и графиков; промежуточные и окончательные расчёты, в том числе расчёт погрешностей полученных результатов.

**Обсуждение результатов:** анализ точности проведённых измерений и достоверности результатов; обсуждение применимости использованных теоретических предположений; сравнение результатов с табличными (справочными) данными или результатами других экспериментов; обсуждение возможных причин ошибок и способов их устранения.



- При необходимости нужно пользоваться *научной* (или *экспоненциальной*) формой записи числа, подбирая наиболее удобные единицы измерения. Например, если длина объекта определена с точностью  $\pm 5$  см и составляет  $l = 123 \pm 5$  см, то вполне допустимы также записи:  $l = 1,23 \pm 0,05$  м, или  $l = (12,3 \pm 0,5) \cdot 10^{-1}$  м, или  $l = (1,23 \pm 0,05) \cdot 10^3$  мм, и т.п. Не вполне корректно было бы написать  $l = 1230 \pm 50$  мм, поскольку такая запись подразумевает превышение точности как в измеренной величине, так и в оценке погрешности.
- Если погрешность физической величины не указана, то по умолчанию подразумевается, что она измерена с точностью до изменения последней значащей цифры на единицу. Например, запись  $l = 1,23$  м эквивалентна  $l = 1,23 \pm 0,01$  м или  $l = 123 \pm 1$  см, но не эквивалентна  $l = 1230$  мм.

При записи *промежуточных результатов* и в промежуточных вычислениях, проводимых вручную, необходимо сохранять одну лишнюю значащую цифру, чтобы избежать ненужных ошибок округления. При вычислениях на калькуляторе необходимо следить, чтобы значащие цифры не вышли за пределы разрядности. То же касается примитивных средств для обработки данных, таких как электронные таблицы. Рекомендуется пользоваться только инженерными/научными калькуляторами, которые не имеют ограничений по разрядности, а также специализированными средствами обработки экспериментальных и статистических данных.

### 4.3.3. Построение графиков

Пусть между двумя величинами  $x$  и  $y$  предполагается некоторая функциональная зависимость. Измеряя пары значений  $(x_i, y_i)$ , получим набор из  $n$  результатов — экспериментальных «точек»

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\},$$

которые изобразим на графике. Каждое измерение  $(x_i, y_i)$  имеет свою погрешность (случайную и/или систематическую)  $\delta x_i$  и  $\delta y_i$ . На графиках погрешности принято изображать в виде «крестов» размером  $\pm \delta x$  по горизонтали и  $\pm \delta y$  по вертикали.

Рассмотрим простейший случай, когда зависимость предполагается линейной:  $y = kx + b$ . Из-за случайных погрешностей при  $n > 2$  будет невозможно провести прямую, проходящую через все экспериментальные точки. Можно, тем не менее, попробовать провести «наилучшую» прямую, проходящую максимально близко ко всем точкам. В математической статистике такую процедуру называют также *линейной регрессией*.

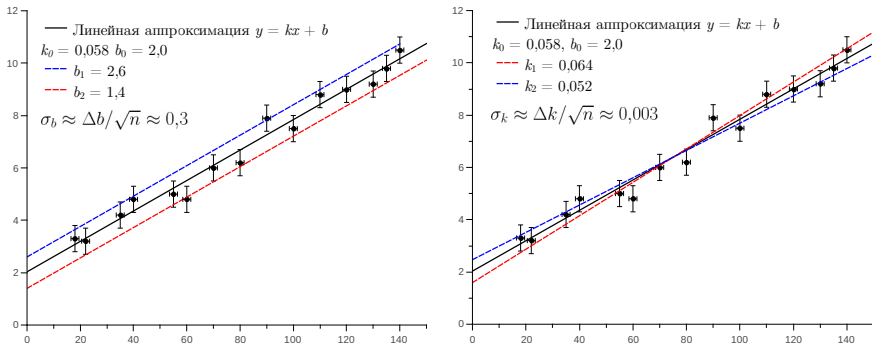


Рис. 4.1. Графический метод проведения прямой и оценки погрешностей

Самый простой и грубый метод — провести наилучшую прямую «от руки». Этот метод, конечно, нестрогий, но весьма наглядный. На практике к нему приходится часто прибегать для грубой и быстрой оценки промежуточных результатов. Для этого нужно приложить прозрачную линейку к графику так, чтобы по возможности кресты всех экспериментальных точек находились максимально близко к проводимой линии, а по обе стороны от неё оказалось примерно одинаковое количество точек.

Построив таким образом «наилучшую» прямую, можно найти её параметры: угловой коэффициент  $k$  и вертикальное смещение  $b$ . Этим же способом можно грубо оценить ошибку определения  $k$  и  $b$ . Сместя линейку вертикально в пределах крестов погрешностей, оценим погрешность  $\delta b$ . Аналогично, изменяя наклон линейки относительно условного «центра масс» экспериментального графика, получим оценку для погрешности углового коэффициента  $\delta k$ . Если известно, что погрешности экспериментальных точек  $(\delta x, \delta y)$  имеют преимущественно случайный характер, результат стоит разделить корень из числа точек:  $\sigma_k \approx \delta k / \sqrt{n}$ ,  $\sigma_b \approx \delta b / \sqrt{n}$  (для систематических погрешностей так делать не стоит).

Эта же процедура позволяет проверить, является ли измеренная зависимость в самом деле линейной: прямая должна пересекать большую часть (хотя бы 2/3) крестов погрешностей. В противном случае можно предполагать существенное отклонение экспериментальной зависимости от линейной теоретической. Отметим, что если кресты погрешностей на графике не отмечены, такой анализ провести затруднительно.

Существуют и аналитические методы подбора параметров (см. гл. 3), минимизирующие отклонения экспериментальных точек от некоторой теоретической зависимости (например, *метод наименьших квадратов*). Студентам первого курса рекомендуется осваивать их постепенно, по мере накопления опыта экспериментальной работы.



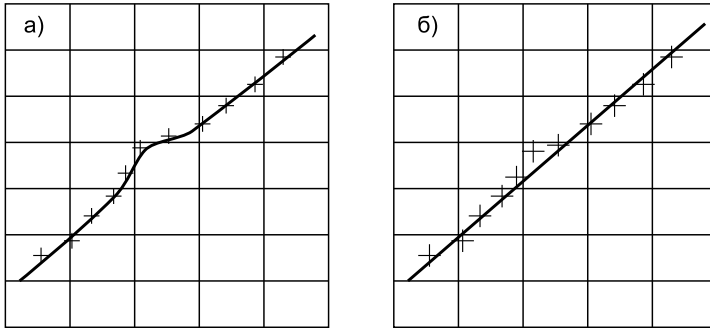


Рис. 4.2

**Пример.** На рис. 4.2 изображены одни и те же экспериментальные точки при разных погрешностях измерений, график 4.2а, несомненно, указывает на нерегулярный ход изучаемой зависимости (кривая линия). Те же данные при больших погрешностях опыта (рис. 4.2б) успешно описываются прямой линией. Без указания крестов погрешностей разделить эти два случая было бы невозможно.

**Нелинейные зависимости.** Если теория предсказывает *нелинейную* функциональную зависимость между величинами, часто можно сделать *замену переменных* так, чтобы результирующий график получался линейным.

Заметим, что аналитические методы позволяют подбирать параметры и для нелинейных зависимостей. Хотя готовых формул для общего случая не существует, задача легко решается численно — и в большинстве современных программ обработки данных это сделать не сложнее, чем построить наилучшую прямую. Построение прямой является наиболее наглядным и позволяет проверить «разумность» полученных результатов, сверив их с построением «от руки».

**Пример.** Высота и время падения груза без начальной скорости в поле тяжести связаны соотношением  $y = \frac{1}{2}gt^2 + y_0$ . Для того, чтобы получить линейную зависимость, можно построить график в координатах  $(y, t^2)$ . По угловому коэффициенту наилучшей прямой можно в таком случае вычислить ускорение свободного падения:  $k = \frac{1}{2}g$ .

**Пример.** В термодинамике и химии часто встречается зависимость вида  $y = Ce^{-ax}$ . Чтобы определить коэффициенты  $C$  и  $a$ , можно построить график в координатах  $(u, v)$ , где  $u = \ln y$  и  $v = \frac{1}{x}$ . В таком случае, как нетрудно видеть,  $u = -av + \ln C$ .

Допускается использования программных пакетов для обработки линейных и нелинейных зависимостей при условии, что студент понимает все детали применяемой процедуры обработки.

#### 4.3.4. Рекомендации по оформлению графиков

Основная цель использования графиков — *наглядность* отображения результатов. В связи с этим к ним предъявляются следующие требования:

- график обязательно должен иметь подпись (заглавие) с кратким описанием его содержания (графики — это первое, на что обращает внимание читатель, еще до прочтения текста отчёта!);
- подписи, данные и линии не должны быть нагромождены друг на друга так, что препятствовало бы их чтению;
- оси на графике должны быть *подписаны*: указаны буквенное обозначение величины и её единицы измерения; если величина безразмерна, указывается «отн. ед.» (относительные единицы);
- на осях должны быть отмечены *масштаб* и *положение нуля*; масштаб обозначается несколькими отметками с подписанными значениями и дополнительными малыми отметками без подписей; масштаб должен быть удобным для чтения (использованы «круглые» числа, делящиеся на 10, 5 или 2);
- масштаб осей и начало отсчёта должны быть выбраны так, чтобы экспериментальные данные занимали всю площадь листа, отведённую под график;
- если график строится не «от нуля», это следует подчеркнуть отдельно, например «разрывом» оси;
- при необходимости сравнения данных из разных серий измерений, их следует размещать на одном графике, обозначая их разными символами или цветами;
- график с несколькими сериями данных должен быть снабжен «легендой», в которой указано соответствие серий данных и их обозначений; экспериментальные «точки» должны изображаться *символами конечных размеров* (позволяющими отличить их от случайных «пятен»);
- точки не должны быть без необходимости соединены линиями; также не нужно подписывать положение каждой точки графика (при необходимости можно указать положение 1-2 *особых* точек, если это не загромождает график);

- все экспериментальные точки должны быть снабжены *крестами погрешностей*, размер которых соответствует инструментальной погрешности измерения соответствующей величины (либо вычисленной по результатам косвенных измерений); кресты погрешностей можно не отмечать, только если погрешности малы (настолько, что они не будут видны на графике) или не известны;
- если теория предполагает некоторую (например, линейную) функциональную зависимость, на график должна быть тонкой линией нанесена соответствующая теоретическая кривая; расчёт параметров этой кривой (например, коэффициентов МНК для линейной зависимости) должен проводиться отдельно в тексте отчёта — с указанием используемых методов и формул; результаты таких расчётов и их погрешности указываются в легенде графика или в подписи к нему;
- оптимальный размер графика — от четверти до половины страницы (при условии, что отчёт оформляется на страницах формата А4).

На рис. 4.3 приведён пример того, как *не надо* строить графики. В нём собраны наиболее типичные ошибки, совершаемые студентами. Предлагаем читателю выявить их самостоятельно. Для сравнения на рис. 4.4 изображён график для *тех же данных*, выполненный с соблюдением изложенных выше указаний.

#### 4.4. Некоторые типичные ошибки обработки данных

**Нахождение углового коэффициента по среднему от частного.** Студент измеряет сопротивление резистора по зависимости  $U(I)$ . Получив некоторое количество экспериментальных точек  $(I_i, U_i)$ , и пользуясь законом Ома  $R = U/I$ , он вычисляет сопротивление для каждого измерения  $R_i = U_i/I_i$ , а затем определяет сопротивление резистора как среднее значение  $R_0 = \langle R_i \rangle = \frac{1}{n} \sum R_i$ . Что не так с этим методом (результат-то получается вполне «разумным»)?

Во-первых, применять процедуру усреднения можно только при повторении *одного и того же* измерения. В данном случае значения  $R_i$  относятся к *разным* измерениям, так как параметры системы каждый раз изменялись. Во-вторых, не была проверена линейность зависимости  $U(I)$ , то есть справедливость закона Ома (ведь существуют и нелинейные элементы, для которых он не выполняется). В-третьих, даже если зависимость можно считать линейной, может оказаться так, что она не проходит через ноль (например, из-за сдвига нуля у вольтметра или амперметра) — тогда формула  $R_i = U_i/I_i$  не годится.

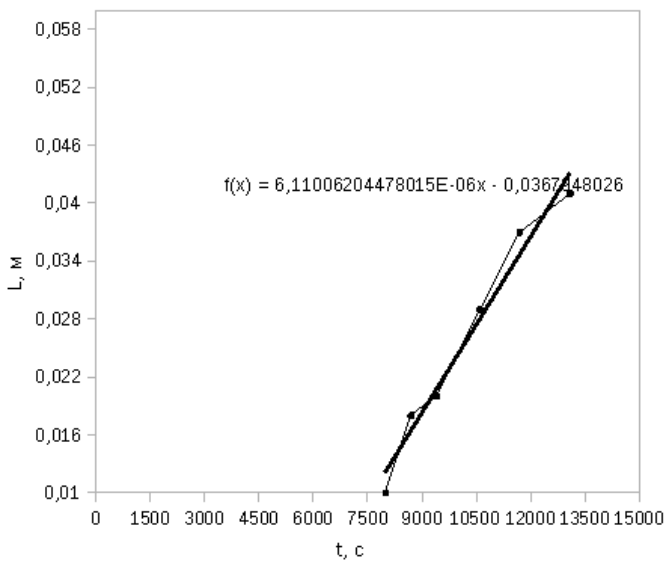


Рис. 4.3. Пример неправильно построенного графика

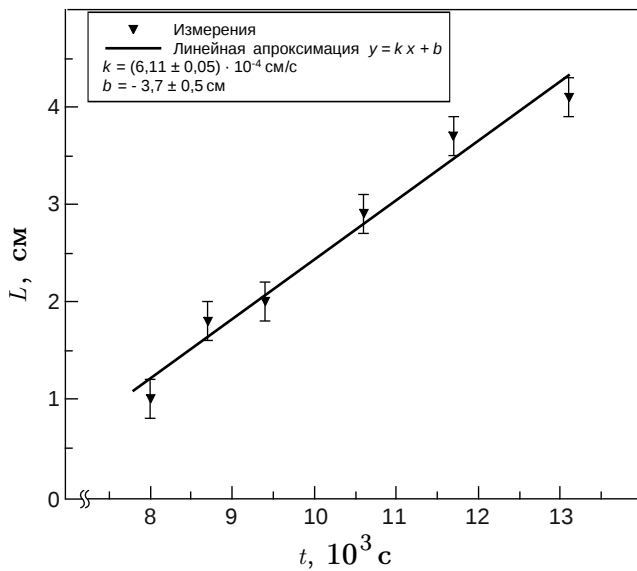


Рис. 4.4. Пример корректного графического представления данных

И наконец, даже если выполнена линейность и зависимость проходит через ноль, вычисление таким способом чревато большими погрешностями. Нетрудно видеть, что среднее значение  $\langle U_i/I_i \rangle$  по сути есть среднее тангенсов углов наклона линий, проведённых из начала координат в экспериментальную точку. Как известно, функция  $\operatorname{tg} x$  при  $x > \pi/4$  очень резко возрастает (и стремится к бесконечности при  $x = \pi/2$ ). В таком случае даже небольшое «шевеление» экспериментальной точки, особенно если она находится достаточно близко к оси ординат, может привести к резкому увеличению вклада этой точки в итоговый результат.

Таким образом, «разумный» результат студента — плод удачного стечения многих обстоятельств. Правильный — обоснованный и надёжный — алгоритм нахождения сопротивления: построить график  $U(I)$ , убедиться в его линейности, и построить наилучшую прямую. Угловый коэффициент этой прямой и будет наилучшей оценкой для сопротивления резистора.

**Недооценка систематической погрешности.** Студент измеряет сопротивление резистора, действуя по правильному алгоритму, описанному выше. При измерениях используются вольтметр и амперметр с классом точности 0,5. Получив большое число экспериментальных точек и построив наилучшую прямую методом наименьших квадратов (формула (3.7)), студент находит сопротивление (например,  $R = 5,555 \text{ Ом}$ ) и его погрешность по формуле (3.10), которая оказывается равна  $\sigma_R = 0,003 \text{ Ом}$ . Окончательный результат измерения записывается как  $R = 5,555 \pm 0,003 \text{ Ом}$ .

Выходит так, что с помощью приборов, относительная погрешность которых составляет 0,5%, получен на порядок более точный результат  $\epsilon \approx 0,05\%$ . Возможно ли такое?

Ситуация эта вполне реальна и встречается в учебной лаборатории довольно часто. Дело в том, что метод наименьших квадратов позволяет оценить *только случайную ошибку* — и она в самом деле может оказаться довольно мала. Однако учебные приборы далеки от совершенства и их ошибка имеет в основном *систематический* характер. Поэтому в данном случае при записи конечного результата необходимо учесть систематическую ошибку, относительная величина которой по составляет не менее 0,5% (согласно классу точности прибора) и значительно превосходит случайную. Результат эксперимента стоило бы записать как  $R = 5,55 \pm 0,03 \text{ Ом}$ .

Всё же, может стать и так, что инструментальные ошибки наших вольтметра и амперметра имеют случайный характер, и мы в самом деле добились кратного повышения точности за счёт много-

кратных повторений измерений (сколько нужно измерений, чтобы увеличить точность на порядок?). Это нетрудно проверить, если заменить вольтметр или амперметр на аналогичный и повторить опыты. Таким образом мы превратим систематическую ошибку *одного* прибора в случайную ошибку *множества* приборов. Если отклонение нового результата от исходного значительно превысит величину  $\pm 0,003$  Ом, гипотеза о случайности инструментальной ошибки будет опровергнута, то есть погрешность отдельного прибора действительно имеет в основном систематический характер.

**Аппроксимация полиномом.** Студент получает набор экспериментальных данных  $\{x_i, y_i\}$ , которые по теории должны ложиться на прямую. Нанеся точки на график, студент видит, что на прямую они ложатся не очень хорошо (см. рис. 4.5а). Для того, чтобы точки лучше ложились на график, студент решает использовать функцию аппроксимации полиномом (например, квадратичным), имеющуюся в наличии во всех электронных таблицах и программах обработки данных. Можно ли так делать?

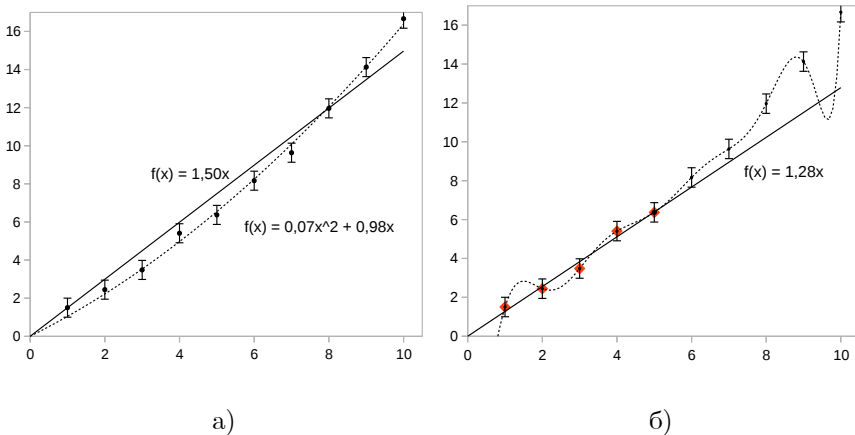


Рис. 4.5. Примеры аппроксимации данных: а) сплошная линия — линейная аппроксимация, пунктир — квадратичная аппроксимация; б) сплошная линия — линейная аппроксимация по нескольким начальным точкам, пунктир — аппроксимация полиномом высокой степени.

Ответ на вопрос зависит от того, какую цель преследует студент. Если цель — добиться того, чтобы «точки лучше ложились на график», то всё сделано правильно. Если говорить «по-научному» — это попытка решить задачу *интерполяции* экспериментальных данных:

по ограниченному набору  $\{x_i, y_i\}$  получить аналитическую функцию  $y = f(x)$ , позволяющую рассчитывать значения  $y$  при произвольном  $x$ . В учебном практикуме это может быть, например, задача построения калибровочного графика. В таком случае можно лишь дать практический совет — стараться использовать для интерполяции полином *как можно меньшей степени* (дело в том, что при слишком большой степени функция почти наверняка станет сильно немонотонной, а это вовсе не то, что хочется иметь в качестве результата интерполяции, см. рис. 4.5б).

Однако ни в одной лабораторной работе курса общей физики задача интерполяции не является основной целью работы! Как правило, цель — проверить теоретические закономерности и измерить физические характеристики системы. Если нет теории, предсказывающей и объясняющей квадратичную зависимость, проделанная студентом процедура *бесполезна*, поскольку вычисленным коэффициентом полинома затруднительно приписать физический смысл.

Правильно было бы в такой ситуации выявить, на каком участке зависимости линейный закон *выполняется*, оценить его границы, и по нему построить наилучшую прямую (сплошная линия на рис. 4.5б). Для участка, отклоняющегося от предсказываемой линейной зависимости, следует теоретически проанализировать причины отклонения и по возможности предложить уточнение теории. Возможно, стоит ожидать не квадратичное, а кубическое отклонение? Различить их на ограниченном наборе данных с большими погрешностями невозможно! Имея достаточное количество точек предложенную теорию можно проверить и лишь после этого аппроксимации более сложной функцией можно придать физический смысл.

## 5. ПРИЛОЖЕНИЕ

### 5.1. Корреляции

Напомним, если величины  $x$  и  $y$  независимы, то среднее значение (математическое ожидание) произведения отклонений  $\Delta x = x - \bar{x}$  и  $\Delta y = y - \bar{y}$  равно нулю:

$$\overline{\Delta x \cdot \Delta y} = \overline{\Delta x} \cdot \overline{\Delta y} = 0.$$

Если же  $x$  и  $y$  не являются полностью независимыми, среднее значение произведения их отклонений может быть использовано как количественная мера их зависимости. Наиболее употребительной мерой зависимости двух случайных величин является *коэффициент линейной корреляции*:

$$r_{xy} = \frac{\langle \Delta x \cdot \Delta y \rangle}{\sigma_x \cdot \sigma_y}. \quad (5.1)$$

Нетрудно проверить (с помощью неравенства Коши–Буняковского), что  $-1 \leq r \leq 1$ . В частности, для полностью независимых величин коэффициент корреляции равен нулю,  $r = 0$ , а для линейно зависимых  $y = kx + b$  нетрудно получить  $r = 1$  при  $k > 0$  и  $r = -1$  при  $k < 0$ . Примеры промежуточных случаев представлены на рис. TODO.

**Замечание.** Угловой коэффициент прямой в задаче линейной регрессии (3.7) выражается через коэффициент корреляции как

$$k = r_{xy} \frac{\sigma_y}{\sigma_x}.$$

Если коэффициент  $r_{xy}$  близок к единице, говорят, что величины *коррелируют* между собой (от *англ.* correlate — находиться в связи).

**Отсутствие корреляции  $\neq$  независимость.** Отметим, что (2.6) — необходимое, но не достаточное условие независимости величин. На рис. TODO приведён пример очевидно зависимых  $x$  и  $y$ , для которых  $r \approx 0$ .

**Корреляция  $\neq$  причинность.** Ещё одна типичная ошибка — исходя из большого коэффициента корреляции ( $r \rightarrow 1$ ) между двумя величинами сделать вывод о функциональной (причинной) связи между  $x$  и  $y$ . Рассмотрим конкретный пример. Между током и напряжением на некотором резисторе имеет место линейная зависимость  $U = IR$ , и коэффициент корреляции  $r_{UI}$  действительно равен единице. Однако *обратное в общем случае неверно*. Например, ток в резисторе коррелирует с его температурой  $T$ ,  $r_{IT} \rightarrow 1$  (больше ток — больше тепловыделение по закону Джоуля–Ленца), однако ясно, что нагрев резистора извне не приведёт к повышению тока в нём (скорее наоборот, так как сопротивление металлов с температурой растёт). Ошибка отождествления корреляции и причинности особенно характерна при исследовании сложных многофакторных систем, например, в медицине, социологии и т.п.

## 5.2. Свойства точечных оценок

Если измеряется одна физическая величина  $x$ , то можно поставить задачу по конечному набору данных  $\mathbf{x} = \{x_i\}$  ( $i = 1 \dots n$ ) оценить параметры случайного распределения, которому подчиняется  $x$ . В частности, найти среднее значение (математическое ожидание)  $\bar{x}$  и дисперсию  $\sigma^2$ .

Если результатом оценки параметра является просто число — без указания интервала, в котором может лежать истинное значение, — такую оценку называют *точечной*. Пример точечных оценок дают формулы для выборочного среднего (1.1):

$$\bar{x} \approx \langle x \rangle = \frac{1}{n} \sum_i x_i \quad (5.2)$$

и выборочной дисперсии (1.2):

$$\sigma^2 \approx s_n^2 = \frac{1}{n} \sum_i (x_i - \langle x \rangle)^2. \quad (5.3)$$



Оценка параметров должна давать правильное значение хотя бы в пределе большого числа измерений. Если при  $n \rightarrow \infty$  оценка стремится к истинному значению параметра,

$$\lim_{n \rightarrow \infty} \hat{\theta}(\mathbf{x}) \rightarrow \bar{\theta},$$

то её называют *состоятельной*. Можно показать (см. [1]), что если у распределения, которому подчиняется случайная величина, существуют конечные средние и дисперсия, то оценки (5.2), (5.3) являются состоятельными.

**Несмещенные оценки.** Рассмотрим случай малого числа измерений ( $n \geq 1$ ). Тогда даже если оценка состоятельна, она может давать довольно большую ошибку. При фиксированном  $n$  функцию оценки  $\hat{\theta}(x_1, \dots, x_n)$  можно рассматривать как случайную величину с некоторым распределением, отличающимся от распределения измеряемой величины. Естественно потребовать, чтобы среднее (математическое ожидание) этого распределения совпадало с истинным значением искомого параметра:

$$\overline{\hat{\theta}(\mathbf{x})} = \bar{\theta}.$$

В таком случае оценку называют *несмещённой*.

Нетрудно показать, что выборочное среднее (5.2) является несмещённой оценкой. А вот оценка  $s_n^2$  из (5.3) таким свойством не обладает. Математическое ожидания для величины  $s_n^2$  при фиксированном  $n$  оказывается равно  $\overline{s_n^2} = \frac{n-1}{n} \sigma^2$  (предлагаем в качестве упражнения проверить данное утверждение самостоятельно). Именно поэтому при малых  $n$  для оценки дисперсии рекомендуется использовать формулу (1.4):

$$\sigma^2 \approx s_{n-1}^2 = \frac{1}{n-1} \sum_i (x_i - \langle x \rangle)^2.$$

**Эффективность оценки.** Для сравнения разных методов оценки очень важным свойством является их *эффективность*. На качественном уровне эффективность — величина, обратная разбросу значений  $\hat{\theta}(\mathbf{x})$  при применении к разным наборам данных  $\mathbf{x}$ . Как обсуждалось выше, оценка  $\hat{\theta}(\mathbf{x})$  есть случайная величина, подчиняющаяся некоторому, в общем случае неизвестному, распределению. Среднее  $\overline{\hat{\theta}(\mathbf{x})}$  по этому распределению определяет смещение оценки. А его дисперсия  $\sigma^2(\hat{\theta})$  — как раз мера ошибки в определении параметра. Выбирая между различными методами, мы, естественно, хотим, чтобы ошибка была минимальной. Разные статистические методы обладают разной эффективностью и в общем случае при конечном  $n$  величина  $\sigma^2(\hat{\theta})$  никогда не будет равна нулю.

Теорема, устанавливающая максимальное значение эффективности оценки, рассмотрена в п. 5.3.

### 5.3. Максимальная эффективность оценки (граница Рао–Крамера)

Максимальная эффективность оценки ограничена теоремой Рао–Крамера.

**Утверждение.** Пусть оценка  $\hat{\theta}$  параметра  $\theta$  является несмещённой, тогда всегда выполняется неравенство:

$$\sigma^2(\hat{\theta}) \geq \frac{1}{I(\theta)}, \quad (5.4)$$

где

$$I(\theta) = \overline{\left(\frac{\partial \ln L}{\partial \theta}\right)^2}. \quad (5.5)$$

Здесь  $L(\mathbf{y}, \theta)$  — введённая в п. 3.1 функция правдоподобия (вероятность получить набор результатов  $\mathbf{y}$  при заданном параметре  $\theta$ ). Функцию  $I(\theta)$  также называют *информацией Фишера*.

**Доказательство в одномерном случае.** Обозначим

$$U \equiv \frac{\partial \ln L}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta}$$

и найдём математическое ожидание этой функции:

$$\bar{U} = \int U \cdot L dy = \int \frac{\partial L}{\partial \theta} dy = \frac{\partial}{\partial \theta} \int L dy = 0.$$

Теперь рассмотрим ковариацию параметра  $\theta$  и функции  $U$ :

$$\overline{\hat{\theta} \cdot U} = \frac{\partial}{\partial \theta} \int \hat{\theta} L dy = \frac{\partial \bar{\theta}}{\partial \theta}. \quad (5.6)$$

Для несмещённых оценок математическое ожидание оценки параметра равно самому значению параметра:  $\bar{\hat{\theta}} = \theta$ , поэтому последнее выражение есть просто единица. Согласно неравенству Коши–Буняковского имеем

$$\sigma^2(\hat{\theta}) \cdot \sigma^2(U) \geq \left| \overline{\hat{\theta} \cdot U} \right| = 1,$$

откуда и следует сделанное утверждение.

**Следствие.** Максимальная эффективность достигается в том случае, если величины  $\hat{\theta}$  и  $U$  являются *коррелируют* друг с другом. Оценка, максимизирующая функцию  $L(\mathbf{y}, \theta)$  (метод максимального правдоподобия), является *состоятельной, несмещённой*, кроме того совпадает с оценкой вида  $U(\mathbf{y}, \hat{\theta}) = 0$ , а значит является *максимально эффективной*.

#### 5.4. Погрешности коэффициентов построения прямой

Проведём подробный вывод для погрешностей коэффициентов наилучшей прямой  $\sigma_k$  и  $\sigma_b$ . Воспользуемся общей формулой (2.11) для погрешности косвенных измерений. Считая, что величины  $x_i$  известны точно, запишем для погрешности углового коэффициента

$$\sigma_k^2 = \sum_i \left( \frac{\partial k}{\partial y_i} \right)^2 \sigma_{y_i}^2.$$

Продифференцируем (3.13) по  $y_i$ :

$$\frac{\partial k}{\partial y_i} = \frac{1}{D_{xx}} \frac{\partial}{\partial y_i} \left( \frac{1}{W} \sum w_i x_i y_i - \langle x \rangle \frac{1}{W} \sum w_i y_i \right) = \frac{w_i (x_i - \langle x \rangle)}{W D_{xx}},$$

где  $D_{xx} = \langle x^2 \rangle - \langle x \rangle^2$ ,  $W = \sum_i \sigma_{y_i}^{-2}$ . Под угловыми скобками здесь понимается выборочное среднее с весами  $w_i = 1/\sigma_{y_i}^2$ . Тогда

$$\sigma_k^2 = \frac{1}{W^2 D_{xx}^2} \sum_i w_i^2 (x_i - \langle x \rangle)^2 \sigma_{y_i}^2.$$

Учитывая, что  $w_i \sigma_{y_i}^2 = 1$ , получим

$$\sigma_k^2 = \frac{1}{W D_{xx}}. \quad (5.7)$$

Аналогично, для погрешности свободного члена имеем

$$\sigma_b^2 = \sum_i \left( \frac{\partial b}{\partial y_i} \right)^2 \sigma_{y_i}^2,$$

где

$$\frac{\partial b}{\partial y_i} = \frac{w_i}{W} + \frac{\partial k}{\partial y_i} \langle x \rangle = \frac{w_i}{W} \left( 1 - \frac{x_i - \langle x \rangle}{\langle x^2 \rangle - \langle x \rangle^2} \langle x \rangle \right) = \frac{w_i}{W} \frac{\langle x^2 \rangle - x_i \langle x \rangle}{D_{xx}}.$$

Отсюда, пользуясь (5.7), приходим к формуле (3.11):

$$\sigma_b^2 = \sigma_k^2 \frac{\left\langle \left( \langle x^2 \rangle - x \langle x \rangle \right)^2 \right\rangle}{D_{xx}} = \sigma_k^2 \langle x^2 \rangle. \quad (5.8)$$

**Случай  $\sigma_y = \text{const}$ .** В частном случае метода наименьших квадратов (п. 3.6.1), формула (5.7) упрощается:

$$\sigma_k^2 = \frac{\sigma_y^2}{n D_{xx}}, \quad \sigma_b^2 = \sigma_k^2 \langle x^2 \rangle. \quad (5.9)$$

Здесь величина  $\sigma_y$  может быть оценена непосредственно из экспериментальных данных:

$$\sigma_y \approx \sqrt{\frac{1}{n-2} \sum_i \Delta y_i^2}, \quad (5.10)$$

где  $n-2$  — число «степеней свободы» для приращений  $\Delta y_i = y_i - (kx_i + b)$  ( $n$  точек за вычетом двух связей (3.13)).

Формул (5.9) и (5.10), вообще говоря, достаточно для вычисления погрешности величины  $k$  по известным экспериментальным точкам. Однако часто их объединяют в одно упрощённое выражение. Для этого преобразуем (5.10) следующим образом: учитывая, что  $\langle y \rangle = k \langle x \rangle + b$ , запишем

$$\Delta y_i = y_i - kx_i - b = (y_i - \langle y \rangle) - k(x_i - \langle x \rangle).$$

Возведём в квадрат, усредним и воспользуемся выражением для  $k$  в форме (3.9):

$$\langle \Delta y^2 \rangle = D_{yy} + k^2 D_{xx} - 2k D_{xy} = D_{yy} - k^2 D_{xx}.$$

Таким образом,

$$\sigma_y = \sqrt{\frac{n}{n-2} (D_{yy} - k^2 D_{xx})},$$

и с помощью (5.9) получаем формулы (3.10), (3.11):

$$\sigma_k = \sqrt{\frac{1}{n-2} \left( \frac{D_{yy}}{D_{xx}} - k^2 \right)}, \quad \sigma_b = \sigma_k \sqrt{\langle x^2 \rangle}.$$

## 5.5. Многопараметрические оценки

Однопараметрические оценки просты для понимания и реализации, но относительно редко встречаются на практике. Даже при оценке параметров линейной зависимости  $y = kx + b$  требуется уже два параметра: наклон  $k$  и смещение  $b$ . Все рассмотренные выше методы нахождения оптимальных параметров работают и в многомерном случае, но поиск экстремума функций (например, максимума функции правдоподобия или минимума суммы квадратов) и интерпретация результатов требуют, как правило, использования численных методов.

### 5.5.1. Двумерный случай

Остановимся подробнее на построении прямой. Пусть некоторым методом получены точечные оценки для наилучших значений  $\hat{k}$  и  $\hat{b}$ . Однако самих значений мало — нас интересует область, в которой могут оказаться параметры  $k$ ,  $b$  с некоторой доверительной вероятностью (например,  $P = 0,68$ ) — двумерная доверительная область.

Предположим для простоты, что оценки параметров имеют нормальное или близкое к нему распределение (это разумное предположение, если результаты получены из большого числа независимых измерений).

Если бы  $k$  и  $b$  были независимы, достаточно было бы найти среднеквадратичные отклонения  $\sigma_k$  и  $\sigma_b$ , как это сделано в п. 5.4: тогда искомая доверительная область значений параметров на плоскости  $(k, b)$  представляла бы собой эллипс, оси которого параллельны координатным (см. рис. 5.1а).

Однако, если взглянуть, к примеру, на рис. 4.2б, иллюстрирующий графический метод построения прямой, можно убедиться, что при варьировании наклона  $k$  обязательно меняется и смещение  $b$ . То есть параметры  $(k, b)$  вообще говоря *скоррелированы*. Количественно отклонения параметров будут характеризоваться ковариационной матрицей:

$$D = \begin{pmatrix} D_{kk} & D_{kb} \\ D_{bk} & D_{bb} \end{pmatrix},$$

где  $D_{kk} = \sigma_k^2$ ,  $D_{bb} = \sigma_b^2$  — дисперсии искомых параметров, а

$$D_{kb} = D_{bk} = \langle (k - \langle k \rangle) \cdot (b - \langle b \rangle) \rangle = \rho_{kb} \sigma_k \sigma_b.$$

Коэффициент  $\rho_{kb}$  называют коэффициентом корреляции и он служит показателем «связанности» параметров. Для полностью независимых параметров он равен нулю, а в случае, если параметры нельзя отличить друг от друга — единице.

По известной теореме линейной алгебры, симметричную матрицу можно привести к диагональному виду поворотом координатных осей. Поэтому доверительная область в таком случае будет представлять собой *наклонный эллипс* (см. рис. 5.1б), а наклон его осей будет определяться коэффициентом корреляции  $r_{kb}$ .

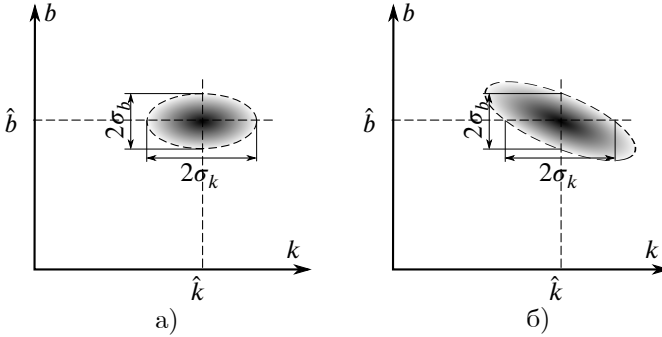


Рис. 5.1. Доверительная область значений коэффициентов прямой а)  $k$  и  $b$  независимы, б)  $k$  и  $b$  скоррелированы.

### 5.5.2. Многомерный случай

Принцип построения доверительной области в многомерном случае точно такой же, как и для одномерных доверительных интервалов. Требуется найти такую область пространства параметров  $\Omega$ , для которой вероятностное содержание для оценки параметра  $\hat{\theta}$  будет равно некоторой наперед заданной величине  $\alpha$ :

$$P(\theta \in \Omega) = \int_{\Omega} L(\mathbf{x}|\theta)d\Omega = \alpha. \tag{5.11}$$

Построение многомерной доверительной области на практике сталкивается с тремя проблемами:

1. Взятие многомерного интеграла от произвольной функции — не тривиальная задача. Даже в случае двух параметров требуется владение методами вычислительной математики. Соответствующие методы реализованы в специализированных программных пакетах.
2. Определение центрального интервала для многомерной гиперобласти является неоднозначной задачей.
3. Даже если удалось получить доверительную область, описать многомерный объект в общем случае непросто, так что представление результатов составляет определенную сложность.

Для решения этих проблем пользуются следующим приемом: согласно центральной предельной теореме, усреднение большого количества одинаково распределенных величин дает нормально распределенную величину. Это же верно и

в многомерном случае. В большинстве случаев, мы ожидаем, что функция правдоподобия будет похожа на многомерное нормальное распределение:

$$L(\theta) = \frac{1}{(2\pi)^{n/2} |D|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T D^{-1}(\mathbf{x}-\bar{\mathbf{x}})}, \quad (5.12)$$

где  $n$  — размерность вектора параметров,  $\bar{\mathbf{x}}$  — вектор наиболее вероятных значений, а  $D$  — ковариационная матрица распределения.

Для многомерного нормального распределения, линии постоянного уровня (то есть поверхности, на которых значение плотности вероятности одинаковые) имеют вид гиперэллипса, определяемого уравнением  $(\mathbf{x} - \bar{\mathbf{x}})^T D^{-1}(\mathbf{x} - \bar{\mathbf{x}}) = \text{const}$ . Для любого вероятностного содержания  $\alpha$  можно подобрать эллипс, который будет удовлетворять условию на вероятностное содержание. Интерес, правда, представляет не сам эллипс (в случае размерности больше двух, его просто невозможно наглядно изобразить), а *ковариационная матрица*. Диагональные элементы этой матрицы являются дисперсиями соответствующих параметров (с учетом всех корреляций параметров).

### 5.5.3. Использование пакета `scipy` для построение оценки

Существует огромное количество программных пакетов для построения численной оценки параметров. Наиболее доступным и широко используемым является пакет `scipy` на языке Python. Приведем здесь только пример вызова процедуры оптимизации.

Пусть есть экспериментальные данные, представленные в виде трех колонок:  $x$ ,  $y$  и  $err$ . Требуется построить наилучшую прямую, описывающую эти данные. Код для этого будет выглядеть следующим образом:

---

```
from scipy.optimize import curve_fit

function = lambda x, a, b: a*x + b

popt, pcov = curve_fit(function, xdata = x, ydata = y, sigma = err)
```

---

После выполнения этого кода, переменная `popt` содержит массив из двух значений, соответствующих оценке  $a$  и  $b$ , а переменная `pcov` содержит ковариационную матрицу для полученных параметров.

Погрешности параметров можно получить как корни из диагональных элементов ковариационной матрицы:

---

```
import numpy as np

sigma_a = np.sqrt(pcov[0,0])
sigma_b = np.sqrt(pcov[1,1])
```

---

**Замечание.** Следует отметить, что существует огромное количество способов оценки оптимальных значений параметров и ковариационной матрицы. Поэтому при использовании того или иного инструмента, всегда следует све-

ряться с документацией и выяснять, что именно он делает. Также следует всегда проверять результаты обработки из качественных, «наивных» сообщений.

## ЛИТЕРАТУРА

- [1] *Тейлор Дж.* Введение в теорию ошибок.
- [2] *Сквейрс Дж.* Практическая физика.
- [3] *Зайдель А.Н.* Погрешности измерений физических величин.
- [4] *Худсон Д.* Статистика для физиков.
- [5] *Идые В., Драйард Д., Джеймс Ф., Рус М., Садуле Б.* Статистические методы в экспериментальной физике.