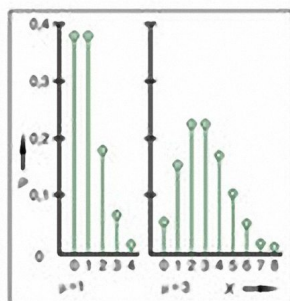


# ХИМИИ

В.И. Дворкин

Метрология  
и обеспечение качества  
химического анализа



ТЕХНОСФЕРА



# М Х и И М и Р и И

В.И. Дворкин

Метрология  
и обеспечение качества  
химического анализа

Издание второе,  
исправленное и дополненное

ТЕХНОСФЕРА  
Москва  
2019



**УДК 543**

**ББК 24.4**

**Д24**

*Научный редактор: Г.Р. Нежиховский, к.т.н.,*

*руководитель сектора ФГУП «ВНИИМ им. Д.И. Менделеева»*

**Д24 Дворкин В.И.**

**Метрология и обеспечение качества химического анализа**

**Издание второе, исправленное и дополненное**

**М.: ТЕХНОСФЕРА, 2019. – 318 с. ISBN 978-5-94836-564-0**

В монографии комплексно рассматриваются теория и практика измерений в аналитической химии. Даются элементы общей теории измерений и методы обработки экспериментальных данных, наиболее важные для аналитической химии, метрология и способы обеспечения качества химического анализа (химическая метрология). Особое внимание уделяется контролю качества химического анализа по конечному результату (внутрилабораторный контроль и межлабораторные эксперименты, включая внешнюю оценку качества (МСИ)).

Книга предназначена для специалистов в области химического анализа, работников лабораторий различного профиля, может использоваться в качестве учебного пособия для студентов и аспирантов химических вузов.

**УДК 543**

**ББК 24.4**

© Дворкин В.И., 2019

© АО «РИЦ «ТЕХНОСФЕРА», оригинал-макет, оформление, 2019

**ISBN 978-5-94836-564-0**

# Содержание

<b>Предисловие</b> .....	7
<b>Предисловие ко II изданию</b> .....	10
<b>Введение</b> .....	11
<b>Глава 1. Элементы общей метрологии</b> .....	15
1.1. Предмет и задачи метрологии. Основные принципы и аксиомы метрологии.....	15
1.2. Величины.....	18
1.2.1. Величина. Размер величины.....	18
1.2.2. Размеры величин и числовые формы.....	19
1.2.3. Размер и значение величины. Измерительное преобразование .....	20
1.2.4. Системы единиц величин. Основные и производные единицы. Размерность.....	21
1.3. Элементы теории измерений.....	23
1.3.1. Классификация измерений .....	23
1.3.2. Принципы, методы и методики измерения .....	26
1.3.3. Средства измерений.....	27
1.3.4. Условия измерений .....	31
1.3.5. Погрешности измерений .....	32
1.4. Эталоны величин. Передача размеров единиц величин.....	33
1.4.1. Эталоны величин.....	33
1.4.2. Передача размеров единиц величин от первичного эталона к образцовым и рабочим средствам измерений .....	34
1.5. Погрешности средств измерений и их нормирование .....	36
1.5.1. Метрологические характеристики средств измерений.....	36
1.5.2. Нормирование метрологических характеристик средств измерений .....	37
1.5.3. Классы точности средств измерений .....	38
1.5.4. Поверка и калибровка средств измерений.....	39
1.6. Российская система измерений.....	41
1.6.2. Утверждение типа средств измерений.....	44
1.6.3. Метрологический надзор .....	45
1.6.4. Государственные метрологические службы [2] .....	46
1.6.5. Международные метрологические организации .....	47
1.7. Количественный химический анализ и особенности его метрологии.....	51
1.7.1. Что измеряют при КХА.....	51
1.7.2. Общий взгляд на метрологию химического анализа .....	52
<b>Глава 2. Элементы математической статистики</b> .....	55
2.1. Общие понятия математической статистики.....	56
2.1.1. Генеральная совокупность и выборка.....	56
2.1.2. Ограниченность возможностей математической статистики	
Статистические модели .....	57

2.1.3. Оценка параметров. Некоторые свойства оценок .....	59
2.1.4. Статистические гипотезы .....	60
2.2. Способы изображения и характеристики выборок .....	62
2.2.1. Изображение выборок .....	63
2.2.2. Статистические характеристики выборок .....	64
2.3. Теоретические распределения .....	68
2.3.1. Непрерывные распределения .....	68
2.3.1.1. Нормальное распределение.....	68
2.3.1.2. Моменты непрерывного распределения .....	69
2.3.1.3. Равномерное и треугольное распределения .....	71
2.3.1.4. Некоторые специальные распределения.....	73
2.3.2. Распределения дискретных случайных величин	
Распределение Пуассона .....	74
2.4. Общие условия применения методов математической статистики и проверка выполнения этих условий .....	76
2.4.1. Проверка гипотезы о распределении генеральной совокупности .....	77
2.4.2. Подозрительно выделяющиеся значения и грубые промахи .....	82
2.5. Некоторые наиболее часто встречающиеся статистические гипотезы и способы их проверки .....	85
2.5.1. Гипотеза о равенстве двух дисперсий .....	85
2.5.2. Сравнение нескольких дисперсий.....	86
2.5.3. Гипотеза о равенстве двух математических ожиданий при независимых (несвязанных) выборках .....	88
2.5.3.1. Случай нормального распределения совокупностей .....	89
2.5.3.2. Непараметрический критерий Вилкоксона – Манна – Уитни .....	93
2.5.4. Гипотеза о равенстве математических ожиданий при связанных выборках .....	96
2.6. Сложение погрешностей и дисперсионный анализ .....	99
2.6.1. Сложение погрешностей. Принцип пренебрежений малыми погрешностями .....	99
2.6.2. Простейший (однофакторный) дисперсионный анализ.....	102
2.6.3. Другие варианты дисперсионного анализа .....	106
2.6.4. Несколько замечаний о планировании эксперимента для последующего дисперсионного анализа .....	108
2.7. Корреляционный и регрессионный анализ .....	108
2.7.1. Проверка взаимосвязи двух величин.....	109
2.7.2. Регрессионный анализ .....	114
2.7.2.1. Графический и некоторые другие методы .....	115
2.7.2.2. Метод наименьших квадратов Гаусса.....	116
2.7.2.3. Случаи невыполнения основных предположений.....	119
2.7.2.4. Анализ остатков. Проверка выполнения предположений и выявление грубых промахов .....	122
2.7.2.5. Многомерная регрессия .....	125

<b>Глава 3. Метрология химического анализа .....</b>	<b>127</b>
3.1. Основные принципы метрологии химического анализа.....	127
3.1.1. Возможна ли корректная передача размера в количественном химическом анализе?.....	128
3.1.2. Особенности методик химического анализа как разновидности методик измерений.....	131
3.2. Образцы сравнения и стандартные образцы.....	135
3.2.1. Погрешность стандартного образца .....	139
3.2.2. Оценка срока годности стандартных образцов.....	141
3.2.3. Оценка неоднородности стандартных образцов.....	144
3.2.4. Способы установления содержания аналитов в стандартных образцах .....	148
3.2.5. Адекватность стандартных образцов.....	151
3.2.6. Утверждение (признание) стандартных образцов.....	156
3.3. Метрологические характеристики методик КХА .....	157
3.3.1. Терминология для прецизионности и ее составляющих .....	157
3.3.2. Терминология для правильности КХА .....	159
3.3.3. Другие метрологические характеристики методик КХА.....	160
3.4. Исследование методик количественного химического анализа .....	160
3.4.1. Градуировочная характеристика и градуировка .....	161
3.4.2. Предел обнаружения .....	163
3.4.3. Оценка прецизионности .....	165
3.4.4. Оценка правильности .....	168
3.4.5. Робастность, селективность и специфичность .....	172
3.4.6. Аттестация (валидация) методик .....	173
3.5. Прослеживаемость и неопределенность результатов анализа.....	174
3.5.1. Прослеживаемость результатов измерений в КХА .....	174
3.5.2. Оценка неопределенности измерений и представление их результатов заказчику .....	177
3.5.2.1. Что такое неопределенность .....	177
3.5.2.2. Причины введения неопределенности и необходимость ее использования .....	180
3.5.2.3. Способы оценки неопределенности результатов КХА .....	182
3.5.2.4. Представление результатов измерений при передаче их заказчику .....	190
3.6. Межлабораторные эксперименты .....	194
3.6.1. Контрольные материалы и дизайн межлабораторного эксперимента .....	194
3.6.2. Способы, применяемые для получения объективной информации о качестве анализа .....	195
3.6.3. Методы обработки результатов межлабораторных экспериментов .....	197
3.6.3.1. Схемы с одним контрольным материалом .....	198
3.6.3.2. Схемы с несколькими контрольными материалами .....	201
3.7. Аккредитация лабораторий .....	206

<b>Глава 4. Обеспечение качества результатов химического анализа</b> .....	210
4.1. О выборе методик количественного химического анализа .....	210
4.2. Отбор пробы .....	212
4.2.1. Отбор пробы вне лаборатории .....	214
4.2.1.1. Основные пути отбора проб .....	214
4.2.1.2. Виды проб .....	215
4.2.2. Отбор пробы в лаборатории .....	217
4.2.3. Сколько пробы нужно отбирать, чтобы она была представительной? .....	219
4.2.4. Оценка и учет погрешности (неопределенности) отбора пробы .....	221
4.3. Статистически контролируемое состояние .....	222
4.4. Проверка приемлемости .....	224
4.5. Внутрिलाбораторный контроль качества .....	227
4.5.1. Контроль стабильности .....	229
4.5.1.1. Планирование процедур контроля стабильности .....	230
4.5.1.2. Период оценки метрологических характеристик .....	237
4.5.1.3. Контроль стабильности с помощью контрольных карт при наличии контрольного материала .....	244
4.5.1.4. Контроль стабильности по контрольным картам в отсутствие контрольного материала .....	259
4.5.2. Другие способы внутрिलाбораторного контроля .....	260
4.5.3. Выбор оптимальных способов ведения внутрिलाбораторного контроля в конкретных ситуациях .....	266
4.6. Внешняя оценка качества результатов количественного химического анализа .....	268
4.7. Компьютеризация при построении системы обеспечения качества аналитических лабораторий [138] .....	272
4.7.1. Для каких целей применяются компьютерные программы в лабораториях .....	272
4.7.2. Основные пути компьютеризации лабораторий .....	274
4.7.3. Достоинства и недостатки разных классов программ .....	275
<b>Приложения</b> .....	282
<b>Литература</b> .....	312

## Предисловие

Химический анализ — прикладной раздел замечательной науки — аналитической химии. Химический анализ — понятие глобальное, оно включает в себя определение химического состава (элементного, молекулярного, изотопного, вещественного и др.) всех веществ и материалов — объектов окружающей среды, минерального и техногенного сырья, металлов, сплавов и изделий из них, объектов машиностроения, судостроения, авиации, космической, атомной и электронной техники, пищевых продуктов, продукции неорганической и органической химии, лекарств, веществ и материалов медицинского назначения, то есть, по существу, всего, из чего состоит окружающий нас природный и рукотворный мир. Свойства веществ и материалов, их качество, их функциональная пригодность неразрывно связаны с их составом, который устанавливают с помощью методов химического анализа.

Требования к составу веществ и материалов с учетом перечисленных сфер их применения непрерывно растут — это одновременное определение большинства элементов Периодической системы с содержаниями от  $10^{-8}$ — $10^{-10}\%$  до 100%, определение структурных составляющих органических соединений, определение изотопов и форм нахождения элементов в материале, установление состава наночастиц, тонких пленок и покрытий, профилей концентрации и многое другое. Все эти задачи решаются методами химического анализа, методами аналитического контроля.

Огромное количество задач аналитического контроля требует использования столь же большого числа аналитических методов, каждый из которых (несмотря на их разнообразие) включает некоторое химическое или физическое воздействие на исследуемый образец, получение специального для данного вещества отклика на данное воздействие в виде аналитического сигнала и измерение последнего. Трудно назвать какие-либо химические реакции, воздействие различных излучений, температурные воздействия, которые бы не применялись в химическом анализе. В этой области знаний используются достижения химии, физики, информатики, приборостроения и других наук.

Важной и неотъемлемой составной частью химического анализа является метрология и обеспечение качества анализа. В конце концов, для пользователя услуг химического анализа не так важны методы, с помощью которых достигнут полученный результат, как его достоверность и точность. В связи с постоянно растущими масштабами аналитического контроля, огромными затратами на его проведение, зависимостью экономики от возможных аналитических ошибок появилась необходимость в создании специального раздела аналитической химии — метрологии химического анализа.

Начало развитию этого раздела положили монографии российского математика В.В. Налимова и немецкого специалиста К. Дерфеля, посвященные применению математической статистики в анализе вещества. Эти книги вышли, соответственно, в 1960 и 1965 годах и практически сразу были востребованы аналитиками всего мира. В последующие годы на теоретической основе, заложенной в этих книгах, начали появляться сначала статьи, а потом монографии и учебники, посвященные метрологии химического анализа. За короткий период метрология химического анализа стала обязательным предметом при подготовке химиков-аналитиков, практически

любая статья и методика в области химического анализа сопровождалась метрологическим разделом. Особо следует отметить приоритет отечественной научной школы в этой области.

Однако после разрушения Советского Союза масштабы исследований в области метрологии химического анализа драматически уменьшились, а центр тяжести работ в этой области переместился за рубеж, где они проводились, главным образом, в рамках создания международных нормативных документов. Возникла потребность в новых учебниках и монографиях по метрологии химического анализа на русском языке.

В 2001 году вышло в свет учебное пособие профессора В.И. Дворкина «Метрология и обеспечение качества количественного химического анализа». Оно сразу нашло своего читателя среди многочисленных химиков-аналитиков, от студентов до преподавателей и работников аналитических лабораторий. За прошедшие годы многое изменилось, появилась новая литература и нормативные документы, и хорошо, что у В.И. Дворкина появилось желание написать новый, современный вариант своей книги. Фактически это новая книга — она по-другому написана, выверены все дефиниции и положения, обобщена информация последних лет. Это актуальное произведение своего жанра.

Не лишая читателя права на познание этой книги, не будем излагать ее содержание, а отметим некоторые особенности. Первым положительным качеством книги является то, что автор книги — аналитик. Ему близки и понятны принципиальные отличия метрологии химического анализа от классических измерений таких физических величин, как масса, длина, время, температура, сила тока и их производных — это отсутствие материального эталона единицы количества вещества (моля), большое число сопутствующих операций идентификации, пробоотбора, пробоподготовки, химических преобразований неизмерительного характера, зависимость измеряемой величины (аналитического сигнала) от состава пробы и др. А именно эти особенности метрологии химического анализа делают ее специфическим разделом метрологии, где определяющими понятиями являются анализ, определение, стандартный образец, методика. Автор книги хорошо знает эти проблемы, умело их излагает, и поэтому читатель-аналитик не чувствует себя в рамках отвлеченной науки, положения которой часто кажутся ему схоластическими.

Книга хорошо структурирована. Она начинается с изложения элементов общей метрологии — терминов и определений, понятий величины, системы единиц. Далее излагаются: теория измерений, принципы, методы и методики измерений, понятия средств измерений, погрешностей измерений, эталонов, передачи размеров единиц, метрологических характеристик, поверки и калибровки. Рассмотрены вопросы метрологического надзора, функционирования отечественных и международных метрологических организаций. В заключение этого раздела автор переходит к количественному химическому анализу как специфическому виду измерений, удачно выделяя и комментируя особенности.

Во второй главе изложены основы математической статистики — теоретического базиса метрологии. Казалось бы, логичнее было бы начать книгу с этого раздела, но автор решил иначе и оказался прав: теория усваивается значительно лучше, когда понятно, зачем она нужна. Отметим, что в главе о математической статистике есть все, что нужно для решения поставленных задач, но ничего лишнего.



В третьей главе автор переходит к основной цели книги — изложению метрологии собственно химического анализа. Здесь всесторонне охарактеризована специфика аналитических измерений, показана возможность использования химически чистого индивидуального вещества в качестве меры. Особое внимание справедливо уделено методикам анализа и стандартным образцам, которые являются главными носителями информации о составе вещества. Далее рассмотрены такие основополагающие понятия, как погрешность, неопределенность, адекватность стандартного образца, метрологические характеристики, аттестация методик, межлабораторные эксперименты и, наконец, аккредитация лабораторий. Этот раздел наиболее ценен для практических аналитиков, студентов и аспирантов, потому что он охватывает все этапы выполнения анализа, совмещая функции монографии, учебника и справочника.

Заключительная глава книги посвящена проблеме обеспечения качества анализа, которая в наименьшей степени раскрыта в научной и учебной литературе для аналитиков. Здесь представлены такие важные этапы аналитических работ, как выбор методики анализа, отбор пробы, проверка приемлемости, оценка метрологических характеристик, контроль стабильности, статистический контроль качества. Завершает раздел по качеству анализа описание внешнего контроля и компьютеризации процесса анализа, включая такой важный вопрос, как лабораторно-информационные системы. В конце книги приведены полезные и хорошо составленные приложения.

Таким образом, В.И. Дворкин написал ожидаемую, хорошую и интересную книгу по актуальным проблемам метрологии и обеспечения качества химического анализа. Эта книга не разочарует своего читателя и потому ей смело можно сказать: «В добрый путь».

*Президент Ассоциации аналитических центров  
«Аналитика», академик Российской академии наук Ю.А. Карпов*

## Предисловие ко II изданию

Первое издание монографии оказалось весьма востребованным и было полностью реализовано, а интерес к книге сохранился — постоянно поступают новые запросы, в связи с чем было принято решение снова издать ее.

Однако с момента выхода в свет первого издания прошло почти 5 лет, и кое-что изменилось. Появились новые термины, на наших глазах меняются определения основных единиц системы СИ и т. д.

В новом издании мы учли все это, исправили некоторые неточности и внесли в текст другие необходимые дополнения.

*Проф. В.И. Дворкин*

## Введение

Рыночные отношения неизбежно сопряжены с ростом требований к качеству продукции и эффективности производства. В химической промышленности это влечет за собой ситуацию, когда управление производственным процессом полностью или почти полностью осуществляется по результатам измерений как «классических» физических величин (давление и температура в реакторах, масса и объем вводимых в них веществ и пр.), так и концентрации веществ в реакционной смеси и конечных продуктах производства и т. д. При этом качество исходной информации, то есть точность измерений, в основном определяет качество продукции. Ужесточение отношений «продавец-покупатель», «производитель-потребитель», перевод этих отношений на формально-правовую основу резко повышает требования к химическому анализу (иногда неправильно называемому «измерением химического состава») и измерениям других характеристик продуктов химического производства, причем особенно важным становится обеспечение сопоставимости результатов измерений в разных лабораториях.

Необходимость защиты прав потребителей пищевых продуктов, различных изделий и материалов также сопряжено с решением чисто метрологических проблем — обеспечением единства измерений в разных лабораториях и высокой точности этих измерений. Здесь речь идет как о химическом анализе, так и о других видах лабораторных исследований. Аналогичные задачи обеспечения и контроля качества результатов химического анализа возникают при исследованиях загрязнения окружающей среды, при анализе объектов биологического и медицинского происхождения, в ходе научных исследований и т. д. Без развития высокоточного химического анализа, требующего метрологического обеспечения, невозможно представить себе геологию, агрохимию, криминалистику, металлургию и т. д.

Таким образом, метрологические аспекты важны в большинстве областей профессиональной деятельности химиков. В связи с этим постоянно развиваются существующие и создаются новые системы обеспечения и контроля качества как химического анализа, так и «классических» измерений. Метрология, смыкаясь с химическими дисциплинами, в основном с аналитической химией, проникает практически во все сферы деятельности химиков и химиков-технологов, а эффективность их профессиональной деятельности все больше зависит от познаний в этой области.

Особую роль метрологические аспекты играют в аналитической химии. Раздел науки, лежащий на стыке *метрологии* и *аналитической химии*, называется «*метрология химического анализа*» [1]. В нем определение количественного содержания компонентов анализируемого вещества — *количественный химический анализ (КХА)* — рассматривается как специфический вид измерений. Почти всегда КХА, выполняемый в лабораториях, является разновидностью косвенных измерений, основанной на находимой экспериментально или известной теоретически зависимости между содержанием компонента и непосредственно измеряемыми величинами (интенсивностью спектральных линий, количеством реагента, взаимодействующего с определяемым компонентом, и т. д.)<sup>1</sup>.

<sup>1</sup> Это чаще всего не относится к простым по устройству автоанализаторам на основе сенсоров, таких как приборы для измерения концентрации СО в полевых условиях — они ближе к «классическим» измерениям.

От других видов измерений КХА отличается принципиально [i]:

- отсутствием эталона единицы количества вещества — моля;
- сложностью и многостадийностью методик КХА, в большинстве случаев включающих, помимо собственно измерений, пробоподготовку (измельчение пробы, ее растворение и т. д.), концентрирование определяемых веществ (экстракция, сорбция) и другие операции;
- наличием во многих методах анализа значительных по величине систематических погрешностей, обусловленных матричными эффектами (влиянием состава конкретного образца на результаты анализа).

Отсутствие эталона моля, связанное с природой этого понятия, делает практически невозможным построение классической системы передачи размера. Для каждого аналита — вещества, концентрация которого измеряется, — необходим свой эталон. Более того, при определении одного и того же аналита в различных объектах используются разные методы анализа, каждый из которых имеет собственные систематические погрешности. Поэтому основой метрологии химического анализа являются многочисленные *стандартные образцы* состава, разные для разных определяемых веществ и объектов анализа. С их помощью осуществляется оценка и контроль погрешностей анализа. Обычным требованием является также использование документированных, тщательно исследованных и валидированных (аттестованных) методик КХА. В ходе исследования методик КХА при их аттестации оценивают целый ряд так называемых метрологических характеристик методики. Это — компоненты прецизионности (характеризующие случайные погрешности), правильность (характеризует систематические погрешности), а также пределы обнаружения определяемых веществ, селективность, специфичность и другие показатели.

Сложность и многостадийность методик КХА, сопряженные с большой долей ручного (хотя и квалифицированного) труда при их выполнении, определяются необходимостью получения адекватного аналитического сигнала от одного или нескольких компонентов сложных объектов, зачастую присутствующих в низких концентрациях. Они влекут за собой потенциальную нестабильность процесса измерений. Поэтому для обеспечения качества результатов КХА используются особые технологии, редкие в классической метрологии<sup>2</sup> — внутрилабораторный контроль и внешняя оценка качества измерений. Основными видами внутрилабораторного контроля являются проверка приемлемости результатов и контроль стабильности анализа с помощью контрольных карт. При проверке приемлемости оценивается качество исследования каждой конкретной пробы. Контроль стабильности позволяет отслеживать и контролировать качество измерений в целом.

Внешняя оценка качества (межлабораторные сравнения или сличения, профессиональное тестирование) есть независимая оценка качества анализа. Лаборатория получает от авторитетной организации пробы с неизвестным ей содержанием аналитов, исследует их и передает результаты анализа для оценки. Результаты

<sup>2</sup> Под классической метрологией мы понимаем метрологию тех измерений, в которых осуществляется передача размера от первичного эталона к образцовым и рабочим средствам измерений — длины, времени, массы, напряжения и т. д., в которых погрешности определяются в основном погрешностями средств измерений, см. гл. 1.

используются для доказательства ее технической компетентности заказчикам и органам по аккредитации, а также как сигнал для улучшения работы лаборатории (при плохих оценках).

Специфика метрологии химического анализа настолько велика, что существует целый ряд национальных и международных документов в этой области.

На русском языке было опубликовано несколько монографий, посвященных методам обработки результатов химического анализа и обеспечению качества химического анализа, например [ii–ix]. Однако они либо посвящены лишь некоторым аспектам проблемы, либо не отражают современную ситуацию. Имеется также книга [x], представляющая собой компиляцию документов международных организаций. Все они не предназначены для преподавания. Исключением является лишь учебное пособие [xi].

Между тем преподавание общей и химической метрологии студентам химических специальностей — требование современности. Метрологические знания необходимы инженерам-химикам и химикам-аналитикам в силу большой роли измерений в их профессиональной деятельности. Однако даже когда химикам и преподается метрология, усвояемость соответствующих знаний, как показывает практика, невысока. Такое положение вещей определяется следующими причинами [xii]:

- отсутствует традиция преподавания химикам метрологии и соответствующие кадры;
- имеющиеся учебники по метрологии ориентированы на студентов машиностроительных вузов. Соответственно, ведется и преподавание этой дисциплины;
- основой метрологии и смежных дисциплин является математическая статистика. Однако преподают ее математики, сами изучавшие главным образом общую теорию вероятности и не представляющие нужды химиков. Сложности создает и специфическое химическое мышление слушателей.

Попыткой решения проблемы стала разработка нами оригинальных курсов «Основы метрологии» и «Метрология химического анализа» для студентов третьего и более старших курсов Московского государственного университета тонкой химической технологии им. М.В. Ломоносова.

Предлагаемая вниманию читателя книга исходно была написана на основе этих курсов и представляет собой систематическое изложение классической метрологии, элементов математической статистики, метрологии и методов обеспечения качества химического анализа. Сведения из «классической» метрологии, как и элементы математической статистики необходимы для понимания остального текста<sup>3</sup>. В контексте данной книги они являются вспомогательными, однако их изложение соответствует требованиям соответствующих разделов образовательного стандарта 200500.62 по направлению «Метрология, стандартизация и сертификация».

<sup>3</sup> Обычная в таких работах трудность — единый термин «анализ» для химического анализа и статистической обработки данных — обойдена следующим способом. Химических анализ обозначается либо просто термином «анализ», либо сочетанием «химический анализ», а статистическая обработка данных — «статистический анализ» либо «анализ» с указанием способа его выполнения («дисперсионный анализ», «корреляционный анализ» и т. д.).

Особое внимание уделено способам обеспечения качества химического анализа «по конечному результату» — внутрилабораторному контролю качества и межлабораторным экспериментам (в том числе внешней оценке качества), которые приобрели в последнее время особое значение. В то же время автор не стремился к исчерпывающему изложению всех вариантов описанных методов, зачастую весьма специфических для разных объектов и методик анализа.

Это практически новая книга. Предыдущая книга автора на ту же тему [xiii] вышла в 2001 году. По сравнению с ней текст полностью переработан с учетом накопившегося у автора опыта преподавания предмета как в системе высшего, так и постдипломного образования. Включены новые разделы, отражающие происходящие в данной сфере изменения, актуализирована терминология, существенно изменившаяся за прошедшее время, учтены изменения, произошедшие в реальной жизни отечественных лабораторий, добавлены многочисленные примеры.

Книга предназначена как для студентов старших курсов и аспирантов химических вузов и химических факультетов университетов, так и для работников лабораторий и всех, кто интересуется общими вопросами аналитической химии.

*Автор выражает особую признательность научному редактору книги, к.т.н. Геннадию Рувимовичу Нежиховскому.*

## Литература к введению

- i. Дворкин В.И. Метрология химического анализа. Большая российская энциклопедия. Т. 20. С. 147–148 (2012).
- ii. Налимов В.В. Применение математической статистики при анализе вещества. — М.: Физматгиз, 1960. — 430 с.
- iii. Шаевич А.Б. Аналитическая служба как система. — М.: Химия, 1981. — 264 с.
- iv. Чарыков А.К. Математическая обработка результатов химического анализа. Методы обнаружения и оценки ошибок. — Ленинград: Химия, 1984. — 168 с.
- v. Шаевич А.Б. Стандартные образцы для аналитических целей. — М.: Химия, 1987. — 184 с.
- vi. Катеман Г., Пийперс Ф.В. Контроль качества химического анализа. — Челябинск: Металлургия, 1989. — 447 с.
- vii. Каплан Б.Я., Филимонов Л.Н., Майоров И.А. Метрология аналитического контроля производства в цветной металлургии. — М.: Металлургия, 1989. — 196 с.
- viii. Буйташ П., Кузьмин Н.М., Лейстнер Л. Обеспечение качества химического анализа. — М.: Наука, 1993. — 188 с.
- ix. Дерффель К. Статистика в аналитической химии. — М.: Мир, 1994. — 268 с.
- x. Причард Э., Барвик В. Контроль качества в химическом анализе. — С.-П.: Профессия, 2011. — 318 с.
- xi. Калмановский В.И. Метрология для химиков. — Нижний Новгород: издатель Ю.А.Николаев, 2007. — 132 с.
- xii. Дворкин В.И. Журнал аналитической химии. 2002. Т. 57. № 2. С. 192–196.
- xiii. Дворкин В.И. Метрология и обеспечение качества химического анализа. — М.: Химия, 2001. — 263 с.

# ГЛАВА I

## ЭЛЕМЕНТЫ ОБЩЕЙ МЕТРОЛОГИИ

### 1.1. Предмет и задачи метрологии. Основные принципы и аксиомы метрологии

*Измерение* — процесс экспериментального получения одного или более значений величины, которые могут быть обоснованно приписаны величине [1]<sup>1</sup>.

При выполнении измерений *всегда* осуществляется сравнение измеряемой величины с другой, подобной ей и принятой за единицу<sup>2</sup>. При этом измеряемую величину *всегда* оценивают в виде некоторого числа принятых для нее единиц, получая *значение величины*.

Измерения различных величин (длины, массы, времени, силы тока, количества вещества, концентрации и т. д.) постоянно сопровождают практическую и научную деятельность человека. Измерения необходимы при изучении природы, поскольку исключительно посредством измерений можно узнать количественные характеристики исследуемых объектов. Можно сказать, что та или иная наука становится точной только тогда, когда благодаря измерениям она получает возможность находить количественные соотношения, выражающие законы природы.

В соответствии с данным выше определением процесс измерения величины в практическом плане представляет собой совокупность операций по применению технического средства, хранящего единицу величины, и заключается в сравнении (в явном или неявном виде) измеряемой величины с ее единицей. Цель этих операций — получение значения величины (или информации о ней) в форме, наиболее удобной для использования.

Скажем здесь, что размер есть свойство, присущее исследуемому объекту (например, концентрация нитрат-иона в растворе), а значение — то, что находится в результате измерения (например, 1,22 ммоль/л); подробнее см. разд. 1.2.3. Так, в простейшем случае, прикладывая линейку с делениями к какой-либо детали, сравнивают ее размер с единицей, хранимой линейкой, и, произведя отсчет, получают значение величины (длины, высоты, толщины и других параметров детали). С помощью измерительного прибора, например вольтметра, сравнивают размер величины, преобразованной в положение указателя, с единицей, хранимой шкалой этого прибора. В измерительном канале измерительной системы также

<sup>1</sup> В [2] измерение определяется как *совокупность операций, выполняемых для определения количественного значения величины*.

<sup>2</sup> Измерение предусматривает описание величины в соответствии с предполагаемым использованием результата измерения, методику измерений и откалиброванную измерительную систему, функционирующую в соответствии с регламентированной методикой измерений и с учетом условий измерений [1]. Подробнее см. ниже.



выполняется сравнение с хранимой единицей, при этом нередко оно происходит в закодированном виде.

Указанную совокупность операций можно назвать измерением, если при этом создан и реализуется ряд условий, а именно:

- возможность выделения измеряемой величины среди других величин;
- установление единицы, необходимой для измерения выделенной величины;
- материализация (воспроизведение или хранение) установленной единицы техническим средством;
- сохранение неизменным размера единицы (в пределах установленной точности) как минимум на срок, необходимый для измерений.

Общими тенденциями развития науки и техники являются рост числа выполняемых измерений, увеличение количества подлежащих измерению величин и расширение диапазонов их измерений, ужесточение требований к точности измерений. Это хорошо видно на примере аналитической химии. Расширение ассортимента выпускаемых промышленностью химических веществ влечет за собой возрастание роли химического анализа на производствах и в области экологического контроля. Из-за растущих требований к составу выпускаемых материалов приходится анализировать самые разные объекты, снижать пределы обнаружения веществ, уменьшать погрешности анализа. Поэтому появляются новые и совершенствуются известные методы анализа, создаются новые приборы и устройства. Все это требует совершенствования метрологического обеспечения этого вида измерений.

Вопросами теории и практики измерений занимается *метрология* (это название происходит от греч. *метрон* – мера, и *логос* – учение, и может быть переведено как «наука об измерениях»). В настоящее время в России принято следующее определение метрологии:

**Метрология** – наука об измерениях, методах и средствах обеспечения их единства и способах достижения требуемой точности<sup>3</sup>.

Как видим, в определении метрологии используются понятия «единство измерений» и «точность измерений».

**Единство измерений** – состояние измерений, при котором их результаты выражены в допущенных к применению в Российской Федерации единицах величин, а показатели точности измерений не выходят за установленные границы [2]. В международной практике прямой аналог этого понятия отсутствует, однако широко применяется тесно связанное с ним понятие «прослеживаемость» (см. разд. 3.5.1).

**Точность измерений** – качество измерений, отражающее близость их результатов к истинному значению измеряемой величины.

Отметим, что на практике единство измерений обеспечивается не всегда, в частности, оно не в полной мере обеспечивается в случае количественного химического анализа (см. разд. 1.7.2 и гл. 3).

Метрология занимает заметное место в ряду других наук. Существуют развитая система научно-исследовательских учреждений метрологического профиля,

<sup>3</sup> Заметим, что международные организации придерживаются другого определения. Так, Международное бюро мер и весов определяет метрологию как «науку об измерениях, охватывающую экспериментальные и теоретические измерения на любом уровне неопределенности в любой области науки и технологии».

многочисленные кафедры метрологии в высших учебных заведениях и т. д. Метролог — массовая профессия. Кроме того, различными видами метрологической деятельности занимаются работники других специальностей. Так, в связи с необходимостью обеспечения качества химического анализа все больше химиков вовлекается в решение метрологических задач.

Выделяют теоретическую, прикладную и законодательную метрологию.

**Теоретическая метрология** создает теоретические основы метрологии. Она решает следующие основные задачи:

- создание и развитие теории измерений и теоретических основ измерительной техники;
- создание и совершенствование теоретических основ построения систем единиц и эталонов;
- разработка теории погрешностей, основанной на математической статистике и теории вероятности;
- разработка общих принципов постановки и проведения измерительного эксперимента;
- разработка теоретических основ вновь возникающих и нестандартно развивающихся видов и областей измерений, таких как измерения на наноуровне (то есть при размерах объектов измерений порядка  $10^{-8}$ – $10^{-9}$  м), измерение ионизирующих излучений, неравновесных процессов и т. д.;
- создание научных основ количественной оценки параметров объектов и технологических процессов, разработка научно обоснованных критериев оценки степени надежности, долговечности и безопасности изделий.

**Прикладная метрология** решает вопросы практического применения в различных сферах деятельности результатов исследований, получаемых в рамках теоретической метрологии, и положений законодательной метрологии. Ее задачами являются:

- создание и совершенствование методов измерений;
- повышение точности измерений;
- разработка новых эталонов, в том числе на новых принципах;
- разработка методов и средств передачи размера единицы от эталона рабочим средствам измерений с минимальной потерей точности;
- обеспечение автоматизации поверочных работ и т. д.

**Законодательная метрология** занимается законодательным обеспечением метрологической деятельности. В большинстве стран, в том числе в России, многие мероприятия по обеспечению единства измерений и требуемой их точности устанавливаются законодательно.

Итогом деятельности законодательной метрологии являются различные документы, имеющие как обязательный (законы), так и рекомендательный (межгосударственные стандарты (ГОСТ), национальные стандарты (ГОСТ Р), межгосударственные рекомендации (РМГ), рекомендации (документы ранга «Р») и т. д.) характер. Заметим здесь, что *термин «стандарт» в метрологии применяется только по отношению к документам, а не к веществам или изделиям.*

Часто тот или иной раздел метрологии называют по отрасли, которую он обслуживает, хотя подобная классификация не вполне строга. Например, (практическую) метрологию в медицине называют «медицинской метрологией», в химии — «хими-

ческой метрологией» и т. д. Настоящая книга в основном посвящена измерениям в химии, то есть химической метрологии.

**Химическая метрология** — раздел метрологии, занимающийся измерениями в химии, главным образом, в количественном химическом анализе<sup>4</sup> (более строгий термин — «**метрология химического анализа**»).

Необходимость выделения метрологии химического анализа в отдельную область обусловлена тем, что измерения в химии (химический анализ) имеют существенные особенности (см. ниже, разд. 1.7.2 и гл. 3).

Как любая точная наука, метрология имеет свои основополагающие принципы. В качестве таких принципов постулируют **три аксиомы метрологии**.

**Аксиома 1. Без априорной информации измерение невозможно.**

Эта аксиома относится к ситуации до измерения и говорит о том, что мы не можем получить оценку интересующего нас свойства, ничего не зная о нем заранее. Отсюда вытекает, что необходимость в измерении вызвана дефицитом количественной информации об изучаемом свойстве объекта и измерение направлено на уменьшение этого дефицита (ясно, что если об этом свойстве известно все, измерять ничего не нужно). Аксиома 1 не говорит, откуда берется эта информация. В реальности для каждого вида измерений априорные знания, необходимые для измерений, накапливаются человечеством в процессе его развития.

**Аксиома 2. Измерение есть не что иное, как сравнение.**

Это констатация того, что единственным способом получения информации о каких-либо размерах является сравнение их между собой. Следствием этой аксиомы, в частности, является введение эталонов величин и системы передачи их размера к образцовым и рабочим средствам измерений. Отсюда же вытекает необходимость использования стандартных образцов состава веществ (см. гл. 3).

**Аксиома 3. Результат измерения без округления является случайным.**

Данная аксиома относится к ситуации после измерения и отражает тот факт, что результат измерения всегда зависит от множества факторов, в том числе и случайных, точный учет которых невозможен в принципе. Отсюда вытекает, что для описания результатов измерений в полной мере необходимо использовать аппарат теории вероятности (главным образом ее ветви — математической статистики).

Аксиома 3 формализует интуитивно понятный ученым с XIX века факт. На самом деле математическая статистика создавалась и развивалась в значительной степени для описания результатов измерений и решения возникающих при измерениях задач.

## 1.2. Величины

### 1.2.1. Величина. Размер величины

Одним из фундаментальных понятий в физике, химии и метрологии является понятие «величина».

**Величина** — свойство явления, тела или вещества, которое может быть выражено количественно в виде числа с указанием отличительного признака как основы для

<sup>4</sup> Количественный химический анализ есть определение содержания (массы, концентрации и т. п.) или количественных соотношений компонентов в анализируемом образце.

сравнения<sup>5</sup>. Это — общее понятие, которое подразделяют, например, на физические, химические и биологические величины.

Близкое определение дано в [3]: «Величина — свойство, общее в качественном отношении многим объектам (системам, их состояниям и происходящим в них процессам), но в количественном отношении индивидуальное для каждого объекта». Типичные физические величины — масса, время, температура, давление; химические — концентрация (массовая, молярная, моляльная) компонента в растворе, количество вещества; биологические — счетная концентрация (число в единице объема) эритроцитов или тромбоцитов в образце крови и т. д.

Из определения величины понятно, что любая величина может проявляться в большей или меньшей степени, т. е. имеет количественную характеристику.

Одно и то же свойство объекта может быть выражено посредством разных величин. Например, степень нагретости тела можно охарактеризовать как температурой, так и средней скоростью движения молекул. Для удобства и обеспечения единства измерений для каждого свойства выбирают одну характеристику, которую узаконивают соглашениями и в дальнейшем ее в основном и используют.

Для того чтобы можно было установить различия в количественном содержании в каждом конкретном объекте свойства, отображаемого физической величиной, вводится понятие *размера величины* (см. также разд. 1.2.3). В реальной жизни вместо «размер» (массы, длины, количества вещества) говорят обычно просто «масса, длина, количество вещества».

### 1.2.2. Размеры величин и числовые формы

Между размерами каждой величины существуют отношения, имеющие ту же логическую структуру, что и отношения между *числовыми формами* (числа разных видов, векторы, матрицы, тензоры). Так, отношения между размерами концентрации или длины, измеренными в одних и тех же единицах, такие же, как между натуральными числами, а между размерами температурного градиента — как между векторами. Поэтому в качестве моделей множеств размеров величин и соотношений между ними используются множества числовых форм с определенными соотношениями. Выбор модели определяется свойствами измеряемых величин.

Числовые формы, приписываемые моделям величин на всем множестве их размеров, условно делят на три группы.

А. Между размерами определены соотношения порядка и эквивалентности  $<$ ,  $>$ ,  $=$ .

**Пример.** К группе А относится твердость, определяемая по шкале Мооса — различные вещества либо тверже, либо мягче друг друга, либо имеют одинаковую твердость в зависимости от того, какой из ряда стандартных минералов они царапают и какой минерал царапает их.

В. Отношения порядка и эквивалентности имеют место как между размерами, так и между разностями в парах размеров.

**Пример.** Таковы время и потенциалы различных полей.

<sup>5</sup> Основой для сравнения может быть единица измерения, методика измерения, стандартный образец или их комбинация.

С. Отношения порядка и эквивалентности имеют место как между размерами, так и между разностями в парах размеров (как в группе В), но, кроме того, определены отношения, подобные сложению и вычитанию, т. е. сумма или разность размеров являются размерами той же величины. К этой группе относится большинство наиболее употребляемых величин.

**Пример.** К группе С относится, например, масса. Действительно, сумма двух масс есть масса тела, уравнивающая их на чашечных весах, а разность — масса тела, которое нужно положить на чашку с более легким телом для приведения весов в равновесие.

Это не так для числовых форм группы В: разность двух моментов времени — не момент времени, разность двух потенциалов — не потенциал, и т. д.

### 1.2.3. Размер и значение величины. Измерительное преобразование

**Понятия «Размер» и «Значение» величины относятся к количественному содержанию в ней свойства, соответствующего понятию величины.** Как сказано выше, присущее исследуемому объекту свойство, которое мы измеряем, называется *размером*.

**Примеры:** температура воды в термостате, давление газа в баллоне, масса навески, количество ионов натрия в некотором объеме воды и т. д.

Размер существует объективно, независимо от того, как мы его измеряем, и не зависит от используемых единиц измерения. Так, водяная баня имеет некоторую температуру — размер температуры этой бани.

Каждому размеру  $Q$  можно приписать положительное число  $q$  — *числовое значение*, которое показывает, во сколько раз этот размер больше *размера* величины  $|Q|$ , принятого за единицу (метра, килограмма, моля и т. д.). Количественное выражение величины

$$Q = q|Q| \quad (1.1)$$

называют *значением физической величины* (например, 1,36 кг, 41 мкмоль и т. д.).

**Измерительное преобразование** — такое преобразование, при котором устанавливается взаимно-однозначное соответствие между размерами двух величин, сохраняющее для некоторого множества размеров преобразуемой величины (называемого *диапазоном преобразования*) все определенные для нее отношения и функции. Преобразование осуществляется с помощью *преобразователя*. Например, при измерении температуры в некотором интервале (*диапазон преобразования*) с помощью термопары (*преобразователь*) она преобразуется в э.д.с.

**Линейное преобразование** — такое измерительное преобразование, при котором результат преобразования  $R$  увеличивается на  $\Delta R$ , если преобразуемая величина  $Q$  увеличивается на  $\Delta Q$ ; если же величина  $R$  увеличивается на  $n\Delta R$ , то результат преобразования  $Q$  увеличивается на  $n\Delta Q$  (при условии, что все величины лежат в диапазоне преобразования).

При измерениях стараются использовать преобразователи, осуществляющие линейное преобразование. Линейность преобразования можно проверить (теоретически и экспериментально) для числовых форм, относящихся к группам В и С (см. разд. 1.2.2).

**Шкала величины** — определенным образом построенная последовательность одноименных величин различного размера. О способах построения шкал величин см. [3].

#### **1.2.4. Системы единиц величин. Основные и производные единицы. Размерность**

В соответствии со второй аксиомой метрологии при измерении измеряемый размер необходимо сравнивать с каким-то другим. Для каждой величины определяют конкретный размер, принимаемый за **единицу величины**. Единицы величин объединяют в системы.

При построении системы величин вначале определяют, какие именно величины будут охвачены этой системой. Затем эти величины делят на *основные и производные*. Для этого прежде всего определяют *число основных величин* следующим образом.

Величины объективно взаимосвязаны. Связи между величинами в общем виде выражаются уравнениями, ниже называемыми **уравнениями величин** (так, скорость есть отношение длины к времени). Определяют число *независимых* уравнений, связывающих входящие в систему величины друг с другом. *Число основных величин (и их единиц) есть разность между полученным числом независимых уравнений величин и числом входящих в них величин*.

Далее определяют, какие именно величины будут основными и каковы будут их единицы. Остальные величины, входящие в систему (и их единицы) будут *производными*. Постулировав единицы основных величин, получают систему уравнений, которую можно решить относительно единиц всех производных величин.

Выбор основных величин и их единиц не может быть сделан абстрактно-теоретически. Их выбирают из соображений эффективности и целесообразности. В качестве основных главным образом выбирают единицы величин, которые могут быть воспроизведены с самой высокой точностью, и от которых технически не слишком сложно передавать размер при воспроизведении единиц производных величин.

Все остальные величины и их единицы (*производные*) образуются с помощью основных величин и их единиц с использованием уравнений величин.

Совокупность выбранных основных и производных величин называется *системой величин*, совокупность их единиц — *системой единиц величин*. Описанный принцип построения систем величин и их единиц был предложен Гауссом в 1832 г.

В ходе развития науки и техники появилось несколько систем величин и их единиц, отличающихся между собой набором входящих в них величин и основными единицами. В настоящее время общепринятой является Международная система единиц (сокращенное обозначение СИ), использования которой, в частности, требуют многие научные и технические журналы. Тем не менее из практических соображений широко используются и внесистемные единицы (баррель, карат, мм. ртутного столба, пункт и т. д.), а в физике — так называемые естественные системы физических величин<sup>6</sup>.

<sup>6</sup> Естественные системы величин используются в теоретической физике. Так, одна из них, предложенная М. Планком в 1906 г., в качестве основных единиц включает постоянную Планка  $\hbar$ , скорость света  $c$ , гравитационную постоянную  $G$  и постоянную Больцмана  $k$ .

Основными преимуществами использования единой системы СИ являются:

- универсальность;
- унификация единиц измерения;
- удобство практического использования единиц, в большинстве случаев лежащих вблизи середины диапазона реально измеряемых величин<sup>7</sup>;
- когерентность (в большинстве основных уравнений величин при использовании единиц системы СИ коэффициенты равны 1);
- простота изучения системы СИ (в частности, в ней разграничены сила и масса).

Наиболее употребляемые величины и их единицы в системе СИ приведены в Приложении 1. В Приложении 2 даны общепринятые множители и приставки к единицам СИ.

В мае 2019 г. система СИ была модифицирована [4]. В частности, переопределены основные единицы и физические константы:

- постоянная Планка  $\hbar$  в точности равна  $6,626070040 \times 10^{-34}$  Дж·с;
- элементарный электрический заряд  $e$  в точности равен  $1,6021766208 \times 10^{-19}$  Кл;
- постоянная Больцмана  $k$  в точности равна  $1,38064852 \times 10^{-23}$ ;
- число Авогадро  $N_A$  в точности равно  $6,022140857 \times 10^{23}$  моль<sup>-1</sup>;
- световая эффективность  $k_{cd}$  монохроматического излучения частотой  $540 \cdot 10^{12}$  Гц в точности равна 683 лм/Вт;
- килограмм остался единицей массы, но его величина устанавливается фиксацией численного значения постоянной Планка равным в точности  $6,626070040 \times 10^{-34}$ , когда она выражена единицей СИ м<sup>2</sup>·кг·с<sup>-1</sup>, что эквивалентно Дж·с;
- ампер остался единицей силы электрического тока; но его величина устанавливается фиксацией численного значения элементарного электрического заряда равным в точности  $1,6021766208 \times 10^{-19}$ , когда он выражен единицей СИ с·А, что эквивалентно Кл;
- кельвин остался единицей термодинамической температуры; но его величина устанавливается фиксацией численного значения постоянной Больцмана равным в точности  $1,38064852 \times 10^{-23}$ , когда она выражена единицей СИ м<sup>-2</sup>·кг·с<sup>-2</sup>·К<sup>-1</sup>, что эквивалентно Дж·К<sup>-1</sup>;
- моль остался единицей количества вещества, но его величина устанавливается фиксацией численного значения постоянной Авогадро равным в точности  $6,022140857 \times 10^{23}$  моль<sup>-1</sup>, когда она выражена единицей СИ моль<sup>-1</sup>;
- кандела, символ кд, является единицей силы света в заданном направлении; ее величина устанавливается фиксацией численного значения световой эффективности монохроматического излучения частотой  $540 \cdot 10^{12}$  Гц равным в точности 683, когда она выражена единицей СИ м<sup>-2</sup>·кг<sup>-1</sup>·с<sup>3</sup>·кд·ср или кд·ср·Вт<sup>-1</sup>, что эквивалентно лм·Вт<sup>-1</sup>.

<sup>7</sup> Конечно, и система СИ в этом смысле не идеальна. Так, единицы давления (паскаль, Па) и электрической емкости (фарада, Ф) далеки от чаще всего измеряемых размеров – атмосферное давление равно  $\sim 10^5$  Па, а 1 фарада близка к электрической емкости Земли.



Формализованным отражением качественного различия величин является их **размерность** (dimention), ее стандартное обозначение —  $\dim$ . Размерность основных величин записывают заглавными латинскими буквами, соответствующими обозначениям величин:  $\dim l = L$  (длина);  $\dim m = M$  (масса);  $\dim t = T$  (время) и т. д. Размерность остальных величин определяют через размерности основных величин по формуле

$$\dim Q = L^{\alpha} \cdot M^{\beta} \cdot T^{\gamma} \cdot \dots, \quad (1.2)$$

где  $L, M, T, \dots$  — размерности основных величин,  $\alpha, \beta, \gamma, \dots$  — показатели размерности, представляющие собой числа (0, целые или дробные), определяемые из уравнений величин (см. выше). При определении размерности производных величин руководствуются следующими правилами.

1. Размерности правой и левой частей уравнения, выражающего связь размерностей двух величин, должны совпадать. Из этого следует, что алгебраическое суммирование допустимо только для величин с одинаковыми размерностями.
2. Алгебра размерностей мультипликативна, т. е. предусматривает только одно действие — умножение. Это очевидно из уравнения (1.2).

Используя эти правила, легко проверить правильность формул, связывающих величины: размерности обеих частей любого уравнения должны совпадать.

Если все показатели размерности равны нулю, то величину называют *безразмерной*. Безразмерные величины бывают относительными (отношение двух величин с одинаковыми размерностями) и логарифмическими (логарифм относительной величины).

**Примеры.** Относительная влажность газов — отношение парциального давления паров воды в газе (в первую очередь, в воздухе) к равновесному давлению насыщенных паров при данной температуре — безразмерная относительная величина. Оптическая плотность растворов  $D = \lg I_0/I$ , где  $I_0$  и  $I$  — соответственно интенсивности падающего и прошедшего через слой раствора света — безразмерная логарифмическая величина.

## 1.3. Элементы теории измерений

### 1.3.1. Классификация измерений

Измерения можно классифицировать разными способами.

**По характеру зависимости измеряемой величины от времени** измерения могут быть *статическими* (измеряемая величина постоянна в течение всего периода измерений) и *динамическими* (изменениями измеряемой величины во время измерений можно пренебречь).

**Пример 1.** Статические измерения — измерение массы сухой колбы на аналитических весах, нахождение концентрации элементов в горных породах рентгенофлуоресцентным методом.

**Пример 2.** Калибровка штриховой меры длины на государственном первичном эталоне. Погрешность (см. разд. 1.3.5) за счет вариации температуры при этом составляет  $\sim 0,01$  мкм/м, что существенно для этой задачи, и такое измерение является динамическим измерений.

**По способу получения результатов** измерения делятся на:

- **прямые**, при которых искомое значение измеряемой величины находят непосредственно из опытных данных. Пример: измерение длины стола рулеткой.
- **косвенные**, когда значение величины находят на основании известной зависимости между этой величиной и величинами, подвергаемыми прямым измерениям. При косвенных измерениях искомую величину рассчитывают по формуле

$$Q = f(X_1, X_2, \dots, X_n), \quad (1.3)$$

где  $Q$  – искомое значение;  $X_1, X_2, \dots, X_n$  – значения величин, измеренных прямым способом;  $f$  – знак функциональной зависимости. Вид функции  $f(X_1, X_2, \dots, X_n)$  либо известен заведомо (например, это какой-либо физический закон), либо его находят экспериментально до выполнения измерения. В последнем случае функцию  $Q = f(X_1, X_2, \dots, X_n)$  называют градуировочной зависимостью, а процесс ее нахождения – построением (нахождением) градуировочной характеристики или просто градуировкой.

**Пример:** измерение концентрации тяжелых металлов (свинца, ртути, кадмия и т. д.) в пищевых продуктах методом атомно-эмиссионного анализа с индукционно-связанной плазмой. Раствор, в который переведена навеска исследуемого образца, распыляется в высокотемпературный факел, в котором происходит испарение частиц. Интенсивность излучения атомов определяемых металлов регистрируется напрямую с помощью матричного детектора. Расчет концентрации (выражаемой в массовых процентах, ммоль/л и т.д.) ведется по градуировочной характеристике, которую строят по результатам нескольким растворам с известной концентрацией определяемых элементов, с учетом массы навески.

- **совместные**, то есть одновременные измерения двух или более различных величин, осуществляемые чаще всего для нахождения зависимости между ними. Пример: градуировка термопары, при которой одновременно измеряют э.д.с. термопары и температуру среды, в которую помещен ее спай.
- **совокупные**, при которых выполняют одновременные измерения нескольких одноименных величин. При этом искомую величину находят, решая систему уравнений, полученных посредством прямых измерений различных сочетаний этих величин.

**Пример** совокупных измерений: нахождение точной массы гирь разновеса 1, 2', 2'', 5, 10 (цифра означает номинальную массу гири в граммах; 2' и 2'' – разные гири с одной и той же номинальной массой 2 г) с помощью образцовой гири массой 1 г, обозначенной как 1<sup>0</sup>. Измерения проводят с комбинациями гирь, взвешивая их на разных чашках весов, показывающих разницу между двумя массами:

$$\begin{aligned} 1 &= 1^0 + a; \\ 1 + 1^0 &= 2' + b; \\ 2' &= 2' + c; \\ 1^0 + 2' + 2'' &= 5 + d; \\ 1^0 + 2' + 2'' + 5 &= 10 + e. \end{aligned}$$

В левой части уравнений – гири, размещаемые на одной чашке весов, в правой – на другой. Величины  $a, b, c, d, e$  – показания весов, т. е. разность массы гирь, помещенных на разные чашки весов. Решая эту систему уравнений, находят массу всех гирь разновеса.



По условиям, определяющим точность измерений, выделяют измерения *максимально возможной точности*, достижимой при существующем уровне техники; *измерения, выполняемые в ходе поверок и калибровок*, — измерения, выполняемые при передаче размера от первичного эталона другим эталонам и рабочим средствам измерений, см. разд. 1.4 и сл.; *технические* измерения, в которых погрешность (см. разд. 1.3.5) результата определяется погрешностью средств измерений.

Измерения максимально возможной точности — это главным образом измерения, выполняемые для воспроизведения единиц величин и для измерения физических констант.

Измерения, выполняемые в ходе поверок и калибровок (см. разд. 1.5.4).

Эти два вида измерений являются *метрологическими* по своему характеру, поскольку проводятся с использованием эталонов и образцовых средств измерения с целью воспроизведения единиц величин и передачи их размера рабочим средствам измерений (см. разд. 1.4.2).

*Технические измерения* — практически все остальные измерения, выполняемые на практике, например, рутинные измерения в заводских и клинико-диагностических лабораториях.

По характеристике точности измерений рассматривают *равноточные* измерения — ряд измерений какой-либо величины, выполненных одинаковыми по точности средствами измерений и в одних и тех же условиях, и *неравноточные* измерения — ряд измерений какой-либо величины, выполненных различными по точности средствами измерений и (или) в разных условиях.

**Примеры.** Измерение сопротивления нескольких резисторов одного номинала одним вольтметром в одних и тех же условиях — равноточные измерения. Взятие навесок вещества на весах различной чувствительности или при различной температуре — неравноточные измерения.

По числу измерений одной и той же величины в ряду измерений последние подразделяют на *однократные* и *многократные*. Однократные измерения выполняют один раз. Часто на практике этого бывает вполне достаточно. При многократном измерении одного и того же размера величины результат получают на основании нескольких следующих друг за другом измерений, т.е. из ряда однократных измерений. За окончательно приводимый результат многократного измерения обычно принимают среднее арифметическое результатов отдельных измерений (см. также разд. 4.4). Формально любое измерение, при котором окончательно приводимый результат рассчитывается по двум и более результатам, многократное. Однако часто (условно) считают измерение многократным, если число отдельных измерений больше или равно 4. В этом случае данные ряда измерений могут быть обработаны методами математической статистики без дополнительных предположений.

**Примеры.** Измерение момента времени по часам или температуры раствора в условиях ее постоянства — однократные измерения. Пятикратное титрование раствора  $\text{KMnO}_4$  в практикуме по аналитической химии в ходе выполнения контрольной работы — многократное измерение.

Кроме того, измерения часто условно делят на *абсолютные*, когда результат выражен в единицах измерения величины, и *относительные*, когда измеряется отношение величины к одноименной величине, играющей роль единицы или при-

нимаемой за исходную. Результаты относительных измерений выражаются либо в долях (*безразмерные величины*), либо в процентах.

**Примеры.** Если результат измерения напряжения выражается в вольтах и равен, например, 224 В — это абсолютное измерение. Если относительная влажность воздуха выражается в процентах и равна, например, 96% — это относительное измерение.

### 1.3.2. Принципы, методы и методики измерения

Основу реализации любого измерения составляет взаимосвязанная триада: принцип, метод и методика измерения.

**Принцип измерения** — совокупность физических явлений, положенных в основу измерения.

**Примеры:** явление поглощения монохроматического излучения лежит в основе спектрофотометрического и атомно-абсорбционного методов измерения концентрации вещества в растворе; эффект силы тяжести составляет принцип измерения массы вещества взвешиванием и т.д.

**Метод измерения** — прием или совокупность приемов сравнения измеряемой физической величины с ее единицей в соответствии с реализованным принципом измерения. Метод измерения зависит от устройства используемых средств измерений. Различают несколько основных методов измерений.

**Метод измерения по определению** заключается в измерении величины в соответствии с определением ее единицы и применяется, как правило, при воспроизведении основных единиц (см. разд. 1.2). Пример: измерения, выполняемые при воспроизведении единицы температуры (кельвина) согласно его определению.

**Метод сравнения с мерой (метод сравнения)** заключается в сравнении измеряемой величины с величиной, воспроизводимой мерой. Пример: взвешивание на рычажных весах с уравниванием гири известной массы.

**Дифференциальный (разностный) метод измерения** заключается в сравнении измеряемой величины с однородной величиной, имеющей известное значение. При этом разность между измеряемой величиной и величиной с известным значением, которую собственно и измеряют, мала по сравнению с самими этими величинами.

**Примеры:** измерения, выполняемые при поверке концевых мер длины сравнением с образцовой мерой на компараторе; спектрофотометрическое определение малых содержаний веществ в анализируемом растворе, когда измеряемая величина — оптическая плотность — представляет собой разницу между оптическими плотностями анализируемого и стандартного (нулевого) растворов.

**Нулевой метод измерения** состоит в том, что результирующий эффект воздействия измеряемой величины и меры на прибор сравнения доводят до нуля.

Пример: измерение количества вещества в газовом хроматографе с помощью детектора по теплопроводности (катарометра). Нулевой метод также реализуется во всех приборах, принцип действия которых основан на измерении электрического сопротивления с помощью моста посредством полного его уравнивания.

В **контактном методе измерения** чувствительный элемент прибора приводится в контакт с объектом измерения.

**Пример:** измерение температуры водяной бани ртутным термометром.

В *бесконтактном методе* измерения чувствительный элемент прибора не приводится в контакт с объектом измерения.

**Пример:** измерение температуры графитовой кюветы пирометром при электротермическом атомно-абсорбционном анализе.

**Методика измерений** — совокупность конкретно описанных операций, выполнение которых обеспечивает получение результатов измерений с установленными показателями точности [6]<sup>8</sup>. Обычно методика измерения регламентируется соответствующим нормативно-техническим документом, в котором излагаются нормы и правила, в соответствии с которыми производятся измерения: требования к выбору средств измерений, процедура подготовки средства измерений к работе, требования к условиям измерений, проведение измерений с указанием их числа, последовательности; способы обработки результатов измерений, включая вычисление и введение поправок и способы выражения точности и так далее, вплоть до требований по технике безопасности (иногда все это называют «унифицированной методикой»). Принято считать, что использование унифицированных методик способствует обеспечению единства измерений.

Особое значение методика имеет для *количественного химического анализа (КХА)*, чаще всего состоящего из множества операций, определяющих результат измерения. Широко используемые методики измерений, в том числе методики КХА, в большинстве случаев проходят *аттестацию*. *Подробнее об исследовании и аттестации методик КХА см. гл. 3.*

### 1.3.3. Средства измерений

**Средства измерений** — технические устройства, предназначенные для измерений, имеющие нормированные метрологические характеристики, воспроизводящие и(или) хранящие единицу физической величины, размер которой принимается неизменным (в пределах установленной погрешности) в течение известного интервала времени. По ряду критериев различают следующие средства измерений.

**По назначению** — образцовые (метрологические) и рабочие.

**Образцовые (метрологические) средства измерений** предназначены для воспроизведения единицы физической величины и(или) ее хранения или для передачи размера единицы другим (как образцовым, так и рабочим) средствам измерений, то есть для *измерений, выполняемых в ходе проверок и калибровок* (см. выше). С их помощью обеспечивается единство измерений в стране. Образцовыми средствами измерений являются эталоны (см. ниже), образцовые приборы, поверочные установки, средства сравнения (компараторы) и др.

**Рабочие средства измерений** предназначены для практических измерений, не связанных с передачей размера единицы величины другим средствам измерений (рабочие измерения). Они позволяют измерять реальные величины и являются самыми многочисленными.

**Примеры:** к рабочим относятся средства измерений, применяемые в лабораториях (весы, рН-метры, вольтметры, манометры), при контроле различных параметров продукции и технологических процессов (датчики, счетчики), торговые весы и т.д.

<sup>8</sup> Закон [2] не делает различий между методом и методикой измерений.

**По уровню стандартизации** — стандартизованные и нестандартизованные. **Стандартизованные средства измерений** изготавливают в рамках требований стандарта. Их технические характеристики соответствуют характеристикам средств измерений аналогичного типа, полученным на основании государственных испытаний, выполняемых при утверждении типа средства измерений (см. разд. 1.6.2). Средства измерений, внесенные в Федеральный информационный фонд по обеспечению единства измерений, чаще всего относятся к числу стандартизованных.

**Примеры** средств измерений указанного типа: вольтметры, весы, мерные колбы, разновесы, стандарт-титры (фиксаналы), широко применяемые в аналитических лабораториях.

**Нестандартизованные средства измерений** предназначены для выполнения специальной измерительной задачи. Они часто являются уникальными, самостоятельно изготовленными. Для того чтобы проведенные с их помощью измерения были достоверными, их точность необходимо оценивать. Такие устройства также могут (но не обязательно) подвергаться государственным испытаниям и вноситься в Федеральный информационный фонд по обеспечению единства измерений.

**По отношению к измеряемой величине** средства измерений делят на основные и вспомогательные.

**Основные средства измерений** производят измерения той величины, значение которой необходимо получить в рамках поставленной измерительной задачи.

**Вспомогательные средства измерений** измеряют ту величину, влияние которой на основное средство измерений или объект измерений необходимо учесть для получения результатов измерений требуемой точности.

**Пример:** при пикнометрическом определении плотности раствора пикнометр является основным, а термометр для измерения температуры воды в термостате — вспомогательным средством измерения.

**По конструктивному исполнению** средства измерений делят на меры, измерительные приборы, измерительные установки, измерительные системы, измерительные комплексы.

**Мера** как средство измерения предназначена для воспроизведения и(или) хранения величины *одного или нескольких заданных размеров*, значения которых выражены в установленных единицах и известны с необходимой точностью.

**Примеры:** гиря; образцовое сопротивление; нормальный элемент Вестона (мера Э.д.с.); кварцевый генератор (мера частоты электрических колебаний); стандартный образец состава вещества и т.д.

Мера выступает в качестве носителя единицы величины и служит основой для измерений. При сравнении с ней размера измеряемой величины получают значение этой величины в тех же единицах.

Меры подразделяют на однозначные, многозначные, наборы мер, магазины мер, установочные. Мера, воспроизводящая величину одного размера, — *однозначная мера*.

**Примеры:** гиря определенной массы, конденсатор постоянной емкости, нормальный элемент Вестона, калибр).

Мера, воспроизводящая величину разных размеров, — *многозначная мера*.

**Пример:** линейка.



Комплект мер разного размера одной и той же величины, необходимый для применения на практике как в отдельности, так и в различных сочетаниях, есть *набор мер*.

**Примеры:** наборы разновесов, калибров и т.д. Наборы мер обычно устанавливаются нормативно-техническими документами.

Набор мер, конструктивно объединенных в единое устройство, в котором имеются приспособления для их соединения в различных комбинациях, называется *магазином мер*.

**Примеры:** магазины электрических сопротивлений, емкостей или индуктивностей.

Мера, предназначенная для приведения показания средства измерения в соответствие с ее известным значением или для контроля неизменности чувствительности средства измерений и приведения его показаний (выходных сигналов) к показаниям, соответствующим чувствительности средства измерений при первичной градуировке, называется *установочной мерой*.

**Пример:** радионуклидный источник, применяемый для контроля стабильности чувствительности радиометрических и дозиметрических приборов.

Близки по назначению к мерам *стандартные образцы*, предназначенные для воспроизведения, хранения и передачи характеристик состава или свойств веществ (материалов), хотя последнее время их выделяют в самостоятельную группу метрологических средств [1, 2]. Подробнее о стандартных образцах состава веществ см. гл. 3.

**Измерительный прибор** — средство измерений, предназначенное для получения значений измеряемой величины в установленном диапазоне. Такой прибор имеет устройство для преобразования измеряемой величины в сигнал измерительной информации и его индикации в доступной для восприятия форме. Почти всегда устройство для индикации имеет шкалу со стрелкой или другим приспособлением, диаграмму с пером или цифровой указатель, с помощью которых можно производить отсчет или регистрацию значений величины. В случае сопряжения прибора с компьютером отсчет может сниматься с дисплея или распечатки.

Важнейшим различием мер и измерительных приборов является число возможных результатов измерений. Если число результатов, которые можно получить данным средством измерений, один или несколько — это мера (однозначная или многозначная), если много — прибор.

**Примеры.** Гиря — однозначная мера, линейка (число возможных результатов несколько десятков) — многозначная мера, а вольтметр (число возможных результатов измерений несколько тысяч) — измерительный прибор.

Имеется несколько классификаций измерительных приборов.

По характеру индикации значений измеряемой величины измерительные приборы разделяют на показывающие и регистрирующие. Первые позволяют только считывать значения измеряемой величины, а вторые — также и регистрировать их. Регистрация показаний может проводиться в аналоговой или числовой форме. Существуют приборы, позволяющие регистрировать одновременно несколько значений одной или нескольких величин.

**Примеры:** микрометр, аналоговый или цифровой вольтметр, часы.



По действию (если таковое производится) измерительные приборы разделяют на интегрирующие и суммирующие. С помощью *интегрирующих измерительных приборов* значение измеряемой величины определяется путем ее интегрирования по другой величине (электрический счетчик электроэнергии, счетчик пройденного расстояния). *Суммирующие измерительные приборы* дают показания, которые функционально связаны с суммой двух или нескольких величин, подводимых по различным измерительным каналам (ваттметр для измерения суммарной мощности нескольких электрических генераторов).

**Измерительные преобразователи** — средства измерений, служащие для выработки сигнала измерительной информации в форме, удобной для передачи, дальнейшего преобразования, обработки и хранения, но не поддающейся непосредственному восприятию наблюдателем. Это — конструктивно обособленные элементы, самостоятельного значения для проведения измерений они, как правило, не имеют. Обычно они являются составными частями более сложных измерительных комплексов и систем автоматического контроля, управления и регулирования.

**Измерительные системы** — совокупность функционально объединенных мер, измерительных приборов, измерительных преобразователей, компьютеров и других технических средств, размещенных в разных точках контролируемого пространства (среды, объекта и т.п.) с целью измерения одной или нескольких величин, свойственных данному пространству (среде, объекту и т.п.). В зависимости от назначения их разделяют на *измерительные информационные системы (ИИС)*, *измерительные контролирующие системы (ИКС)*, *измерительные управляющие системы (ИУС)* и др. Первая из указанных систем представляет измерительную информацию в виде, необходимом потребителю. Вторая — предназначена для непрерывного контроля параметров технологического процесса, явления, движущегося объекта или его состояния. ИУС обеспечивает автоматическое управление технологическим процессом, производством, движущимся объектом и т.п. Эта система содержит элементы сопоставления параметров измерительной информации с нормативными, а также элементы обратной связи, которые дают возможность подводить к номинальным значениям параметры процесса или движущегося объекта, подлежащего управлению. В зависимости от числа измерительных каналов, измерительные системы могут быть одно-, двух- трех- и более канальными. Если система имеет автоматические средства для получения и обработки измерительной информации, то ее называют автоматической измерительной системой. Систему, перестраиваемую в зависимости от цели измерительной задачи, называют гибкой измерительной системой.

**Измерительные комплексы** — функционально объединенная совокупность средств измерений и вспомогательных устройств, предназначенная для выполнения в составе ИИС конкретной измерительной задачи.

**Пример:** измерительные комплексы для оценки качества изготовленных интегральных схем.

*В дальнейшем мы будем рассматривать только однозначные меры и измерительные приборы.*

**По уровню автоматизации** средства измерений делят на неавтоматические средства измерений, автоматизированные средства измерений, автоматические средства измерений.

*Неавтоматическое средство измерений* не имеет устройств для автоматического выполнения измерений и обработки их результатов.

**Примеры:** рулетка, теодолит, пирометр, индикаторная бумага.

*Автоматизированное средство измерений* производит в автоматическом режиме одну или несколько измерительных операций.

**Пример:** компьютеризированный атомно-абсорбционный спектрофотометр с автоматическим пробоотборником.

*Автоматическое средство измерений* производит в автоматическом режиме измерений и все операции, связанные с получением и обработкой результатов измерений, их регистрацией, передачей данных или выработкой управляющего сигнала.

**Пример** — автоматический хроматограф, установленный на химическом реакторе и периодически передающий результаты измерений на пульт управления производством.

### 1.3.4. Условия измерений

Измерения проводят в условиях, при которых все значения влияющих величин поддерживаются в пределах, не выходящих за границы их номинальных значений. Такие условия называют *нормальными*. Они устанавливаются в нормативно-технических документах на средства измерений конкретного вида или при их поверке. При большинстве измерений нормируется нормальное значение температуры (в одних случаях это 20 °С, или 293 К, в других 23 °С, или 296 К). Так, для выпускаемых промышленностью нормальных гальванических элементов нормированное значение их э.д.с. относится к температуре 20 °С. На нормальное значение обычно рассчитана основная погрешность средства измерений. К нормальным условиям часто приводятся результаты измерений, выполненных в других условиях.

**Пример:** пересчет плотности нефти, измеренной при различных температурах и давлениях, к стандартным условиям при товарно-коммерческих операциях [7].

Область значений влияющей величины, в пределах которой изменением результата измерений под ее воздействием можно пренебречь в соответствии с установленными нормами точности, называется *нормальной областью значений влияющей величины (нормальной областью)*.

**Пример.** Для одного из имеющихся в лаборатории автора спектрофотометров нормальная область значений температуры окружающего воздуха составляет  $(20 \pm 5)$  °С.

Область значений влияющей величины, в пределах которой нормируют дополнительную погрешность или изменение показаний средства измерений, называется *рабочей областью значений влияющей величины (рабочей областью)*.

**Пример.** При большой разнице температур шарика жидкостного ртутного термометра и его капилляра вносят «поправку на выступающий столбик» — учитывают дополнительную погрешность.

Условия измерений, в которых измеряемая и влияющая величины принимают экстремальные значения и которые средство измерений еще может выдержать без разрушений и ухудшения его метрологических характеристик, называют *предельными условиями измерения*.

Условия, в которых проводятся измерения с помощью средств измерений, подлежат *нормированию* (см. разд. 1.5.2).

### 1.3.5. Погрешности измерений

Одной из основных метрологических характеристик результатов измерений является погрешность.

**Погрешность измерения** — отклонение результатов измерения от истинного значений измеряемой величины ( $X_i - \tau$ ), где  $X_i$  — результат измерения,  $\tau$  — опорное значение. В теоретических работах часто используется альтернативное определение: погрешность есть *разность между измеренным значением величины и истинным значением величины*. Поскольку вместо неизвестного истинного значения на практике используют опорное значение, оба определения практически эквивалентны.

Погрешность возникает из-за несовершенства процесса измерений.

Конкретные причины и характер проявления погрешностей весьма разнообразны. Соответственно, их классифицируют по многим критериям.

**По характеру проявления** — систематические и случайные погрешности.

**Систематическая погрешность измерения** — составляющая погрешности измерения, остающаяся постоянной или закономерно изменяющаяся при повторных измерениях одной и той же физической величины. В зависимости от характера изменения систематические погрешности подразделяют на постоянные, пропорциональные, периодические, погрешности, изменяющиеся по сложному закону и т.д. Выделяют следующие виды систематических погрешностей.

**Постоянные погрешности** не зависят от измеряемой величины и длительное время (по крайней мере в течение всего периода выполнения измерений) сохраняют свое значение. Они встречаются наиболее часто.

**Пример:** постоянное, отличное от нуля значение холостого опыта в химическом анализе.

**Пропорциональные погрешности** изменяются пропорционально значению измеряемой величины.

**Периодические погрешности** являются периодической функцией времени.

**Погрешности, изменяющиеся по сложному закону**, представляют собой результат совместного действия нескольких видов систематических погрешностей.

**В зависимости от причин возникновения** систематические погрешности подразделяют на инструментальные, погрешности метода измерений, субъективные, погрешности вследствие несоблюдения установленных условий измерений.

**Инструментальные (аппаратурные) погрешности измерений** обусловлены погрешностями применяемого средства измерения. Они возникают из-за износа деталей и прибора в целом, излишнего трения в механизме прибора, неточного нанесения штрихов при калибровке, вследствие несоответствия действительного и номинального значений меры и т.д. В последние годы в этот вид погрешности стали включать также и случайную составляющую погрешности (см. ниже), приписываемую средству измерения.

**Погрешности метода измерений** (теоретические) обусловлены несовершенством принятого метода измерений. Они являются следствием упрощенных представлений о явлениях и эффектах, лежащих в основе измерений.

*Субъективные погрешности измерений* (личные, личная разность) вызваны индивидуальными особенностями оператора.

*Погрешности измерений из-за изменения условия измерений* возникают вследствие неучтенного или недостаточно учтенного воздействия той или иной влияющей величины (температура, давление, влажность воздуха, напряженность магнитного поля, вибрации и др.), неправильной установки средств измерений и других факторов, связанных с условиями измерений.

*Случайная погрешность измерения* — составляющая погрешности измерения, изменяющаяся случайным образом (по знаку и значению) при повторных измерениях одной и той же величины. Случайные погрешности неизбежны и неустранимы и всегда присутствуют в результатах измерений. Они вызывают рассеяние числовых значений измеряемой величины (различие их в последних значащих цифрах) при многократном и достаточно точном ее измерении при неизменных условиях.

Кроме того, выделяют *грубую погрешность измерения (грубые промахи или выбросы)* — погрешность, существенно превышающую ожидаемую при данных условиях проведения измерений.

По способу выражения погрешности делят на абсолютные и относительные.

*Абсолютная погрешность измерения* — погрешность измерения, выраженная в единицах измеряемой величины.

*Относительная погрешность измерения* — отношение абсолютной погрешности измерения к истинному значению. Поскольку в качестве замены неизвестного истинного значения используется опорное значение, на практике относительную погрешность обычно рассчитывают как отношение абсолютной погрешности к опорному значению.

По условиям измерения [измеряемой величины] погрешности делят на статические и динамические. *Статические погрешности измерений* отвечают условиям статических измерений, *динамические* — условиям динамических измерений.

В зависимости от условий измерений рассматривают также *основные и дополнительные погрешности* (см. разд. 1.3.4).

В последнее время наряду с погрешностью для характеристики результатов измерений все чаще рассматривают *неопределенность* (см. разд. 3.5.2 и сл.).

## 1.4. Эталоны величин. Передача размеров единиц величин

### 1.4.1. Эталоны величин

Для обеспечения единства измерений необходимым условием является тождественность единиц, в которых проградуированы все средства измерений одной и той же величины. Это достигается путем точного воспроизведения и хранения установленных единиц величин и передачи их размеров средствам измерений посредством *эталонов единиц величин*.

*Эталон единицы величины* — техническое средство, предназначенное для воспроизведения, хранения и передачи единицы величины [2]. Эталоны единиц, величины,

находящиеся в федеральной собственности РФ, называются *государственными эталонами единиц величин*. Если эталон воспроизводит единицу физической величины с наивысшей в стране точностью, он называется *первичным*. Первичные эталоны основных единиц воспроизводят эти единицы в соответствии с их определением (с другой стороны, то, как определяется единица физической величины, в той или иной степени обусловлено устройством первичного эталона). Первичные эталоны являются государственными.

Эталоны могут представлять собой меры, то есть иметь фиксированный размер.

**Примеры:** эталоны длины и частоты — установки, воспроизводящие метр и герц по их определению.

Однако во многих случаях эталоны воспроизводят единицу в определенном диапазоне величин.

**Пример:** государственный первичный эталон температуры, воспроизводящий Международную температурную шкалу (МТШ-90) в диапазоне 0–961,78 °C (область контактной термометрии) на основе основных реперных точек с интерполяцией между ними эталонным платиновым термометром сопротивления.

При утверждении государственного эталона утверждается государственная поверочная схема, описывающая, как именно используется эталон и как его размер передается другим средствам измерений. На утвержденный в установленном порядке эталон выдается свидетельство, в котором указываются его метрологические параметры и разряд по национальной поверочной схеме. За состояние эталонов, в том числе за их поверку или калибровку, отвечают их владельцы. Государственным первичным эталонам придается столь большое значение, что для каждого из них вводится специальная должность «ученый-хранитель государственного первичного эталона».

Для повышения их точности эталоны единиц постоянно совершенствуются. Это может происходить как за счет постепенного внесения улучшений в их конструкцию, так и при изменении определения единицы.

**Пример.** Вот уже более 100 лет эталон массы (1 кг) — это гиря, хранящаяся в сейфе в Севре, Франция, и имеется риск ее повреждения либо утраты. Кроме того, по-видимому, ее масса несколько изменилась с момента создания в 1889 г. (на  $\sim 0,04$  мкг), что фактически привело к изменению единицы массы. С мая 2019 г. определение килограмма изменится — он будет «привязан» к постоянной Планка и скорости света в вакууме [8–10] (см. также разд. 1.2.4).

#### **1.4.2. Передача размеров единиц величин от первичного эталона к образцовым и рабочим средствам измерений**

Такая передача осуществляется посредством *образцовых средств измерений*. Собственно передача размера от более точного средства измерений к менее точному осуществляется в ходе процедур поверки и калибровки<sup>9</sup>.

*Образцовые средства измерений* представляют собой эталоны, другие образцовые меры, измерительные устройства или измерительные преобразователи, предназначенные для поверки и градуировки по ним других средств измерений.

<sup>9</sup> Подробнее о поверке и калибровке см. ниже, разд. 1.5.4.



Рис. 1.1. Схема передачи размера от первичного эталона к другим средствам измерений

При этом в настоящее время *все* образцовые средства измерений в законодательной метрологии относятся к категории «эталоны».

В самом общем виде метрологическая цепь передачи размеров единиц физических величин показана на рис. 1.1. Она имеет строгую иерархию, передача размеров идет сверху вниз: от первичного эталона к рабочим эталонам; от рабочих эталонов к эталонам 1-го разряда и к рабочим средствам измерений (мерам и измерительным приборам) наивысшей точности; от эталонов 1-го разряда к эталонам 2-го разряда и к рабочим средствам измерений высшей точности и т.д.<sup>10</sup> Нижележащие эталоны и другие средства измерений поверяются или калибруются по находящимся на ступень выше.

В соответствии с принципом пренебрежения малыми погрешностями (см. разд. 2.6.1) погрешность (неопределенность), обеспечиваемая образцовым средством измерений, лежащим выше в цепи передачи размера, должна быть не менее чем в три раза меньше погрешности нижележащего. Реально соотношение этих погрешностей колеблется от 1:2 до 1:8 для разных ситуаций.

Показанная на рис. 1.1 цепь передачи размера используется лишь для нескольких величин, таких как длина, частота, масса и т.д. В других случаях число ступеней в иерархии может быть существенно меньшим. Это число и порядок передачи размера для каждой величины фиксируются в поверочных схемах [11].

**Пример.** В цепи передачи размера кельвина к рабочим контактными термометрам, кроме первичного и рабочего эталонов, присутствуют только эталоны (образцовые термометры) 1-го и 2-го разрядов [12].

Отметим, что во многих случаях корректная цепь передачи размеров вообще не может быть построена. Так обстоит дело, в частности, с количественным химическим анализом (подробнее об этом см. разд. 1.7 и гл. 3).

<sup>10</sup> Эталоны, которым присвоен 1, 2, 3 или 4-й разряд, часто называю «разрядными».

## 1.5. Погрешности средств измерений и их нормирование

### 1.5.1. Метрологические характеристики средств измерений

Метрологическими характеристиками средств измерений называют их технические характеристики, влияющие на результаты и погрешности измерений. Для каждого средства измерений комплекс этих характеристик выбирается и нормируется таким образом, чтобы с их помощью можно было оценить возможную погрешность (неопределенность) результатов измерений.

Основными метрологическими характеристиками средств измерений являются следующие [13].

*Погрешность* средства измерений есть разность между показанием средства измерений и истинным (действительным, опорным) значением измеряемой величины. Это важнейшая характеристика средства измерения.

Погрешности средств измерений, как и погрешности измерений вообще (см. разд. 1.3.5), делят на:

- *статические* и *динамические* (мы рассматриваем только статические погрешности);
- *систематические* и *случайные*. В отличие от случайных систематические погрешности являются функцией измеряемой величины и времени.

Кроме того, при анализе погрешностей средств измерений (компонент погрешностей) условно выделяют *пропорциональные* (измеряемой величине) и *постоянные* (не зависящие от измеряемой величины) погрешности.

В соответствии с определениями, данными в разд. 1.3.5, различают *абсолютную* и *относительную* погрешности, которые принято записывать для основных видов средств измерений следующим образом.

*Абсолютная погрешность  $\Delta$  для меры* есть разность между ее номинальным и действительным значениями:

$$\Delta = X_n - X_d. \quad (1.3)$$

Номинальное значение меры  $X_n$  есть значение величины, приписанное мере или партии мер при изготовлении (например, гиря с номинальным значением 1,000 г). Действительное значение меры  $X_d$  — это величина, приписанная мере на основании ее калибровки или поверки.

**Пример.** В ходе поверки гири с номинальным значением 1,0000 г было установлено действительное значение 1,0022 г. Следовательно, ее абсолютная погрешность равна  $\Delta = 1,0000 - 1,0022 = -0,0022$  г.

*Абсолютная погрешность  $\Delta$  для измерительного прибора* есть разность между его показанием  $X_n$  и действительным значением измеряемой величины  $X_d$ :

$$\Delta = X_n - X_d. \quad (1.4)$$

Действительное значение измеряемой величины  $X_d$  есть значение, найденное при передаче размера (см. выше) и содержащее столь малую погрешность, что оно может быть принято за истинное значение.

Относительная погрешность  $\delta$  средства измерений представляет собой отношение абсолютной погрешности величины  $X(\Delta X)$  к действительному значению  $X_d$ , выраженное в процентах:

$$\delta = (\Delta X/X_d)100\%. \quad (1.5)$$

Поскольку почти всегда  $\delta \ll 1$ , для приборов полагают  $X_n \approx X_d$ :

$$\delta \approx (\Delta X/X_n)100\%. \quad (1.6)$$

**Градуировочная характеристика (статическая характеристика преобразования, функция преобразования)** представляет собой зависимость вида  $Y = f(X)$  выходного сигнала  $Y$  от входного сигнала  $X$ . Она задается (нормируется) для измерительных приборов в форме уравнения, графика или таблицы и официально приписывается данному средству измерений во всем диапазоне измерений. Величину  $f'(X) = dY/dX$  называют **чувствительностью** характеристики преобразования. Часто говорят о чувствительности средств измерений, подразумевая чувствительность соответствующей статической характеристики преобразования. Статическую характеристику преобразования вида  $Y = KX$  называют **линейной**. Для линейной характеристики чувствительность, очевидно, равна  $K$ .

**Цена деления (для шкальных приборов)** — изменение измеряемой величины, которому соответствует перемещение указателя на одно деление шкалы. Для **цифровых приборов** роль цены деления играет цена единицы младшего разряда числа в показании прибора. В случае, когда чувствительность для данного прибора постоянна в каждой точке диапазона измерений, шкалу называют **равномерной**.

Существуют и другие метрологические характеристики, специфические для отдельных видов измерений.

### 1.5.2. Нормирование метрологических характеристик средств измерений

**Нормирование** — установление границ на допустимые отклонения реальных метрологических характеристик средств измерений от их номинальных значений. Нормы на значения метрологических характеристик устанавливаются соответствующими документами.

Реальные значения метрологических характеристик средств измерений определяются при их изготовлении и затем периодически проверяются во время эксплуатации (в ходе калибровок, поверок). Если при этом хотя бы одна из метрологических характеристик выходит за установленные границы, то такое средство измерений либо регулируют, либо изымают из обращения.

Заметим, что нормируются как метрологические характеристики средств измерений, так и условия, в которых они эксплуатируются (**условия применения**), например, температура или давление атмосферного воздуха (см. также разд. 1.3.4). Могут нормироваться погрешности как в нормальной, так и в рабочей области значений величин, влияющих на результаты измерений.

Погрешность средства измерения  $\Delta_{\text{сумм}}$  в нормальных условиях называется **основной погрешностью** и нормируется заданием предела  $\Delta_d$ . В некоторых случаях отдельно нормируются систематическая  $\Delta_c$  и случайная  $\Delta$  составляющие погрешности.



### 1.5.3. Классы точности средств измерений

Для облегчения стандартизации средств измерений, упрощения их выбора для измерений с требуемой точностью и расчета суммарной погрешности многим средствам измерений приписывают *класс точности*.

**Класс точности** — обобщенная характеристика средства измерения, определяемая пределами допускаемых основных и дополнительных погрешностей, а также другими свойствами средства измерений, влияющими на точность осуществляемых с их помощью измерений. Упрощая, можно сказать, что класс точности «включает» все виды погрешностей, характерные для данного средства измерений при правильном его использовании.

Классы точности средств измерений устанавливаются для средств измерений, для которых:

- систематические и случайные погрешности не нормируются отдельно;
- динамическая погрешность пренебрежимо мала.

Способ обозначения класса точности средства измерений определяется способом задания пределов допускаемой основной погрешности. Обычно для этого используется либо *приведенная*, либо *относительная* погрешность.

**Приведенная основная погрешность**  $\gamma$  используется для приборов в случаях, когда их диапазон измерения охватывает и нулевое значение. Она равна отношению абсолютной погрешности прибора к *нормирующему значению*  $X_N$  (это может быть диапазон измерений, длина шкалы и т.д.), которое выбирается для каждого средства измерений по правилам, изложенным, например, в [14], и обязательно приводится в документации на средство измерений.

Пределы допускаемой приведенной основной погрешности нормируются как

$$\gamma = \Delta/X_N = \pm p. \quad (1.7)$$

Здесь где  $p$  — число из ряда

$$1 \cdot 10^n, 1,5 \cdot 10^n, 2 \cdot 10^n, 2,5 \cdot 10^n, 3 \cdot 10^n, 4 \cdot 10^n, 5 \cdot 10^n, \dots \quad (n = \dots, 2, 1, 0, -1, -2, \dots). \quad (1.8)$$

Если точность средства измерений нормируется по формуле (1.7), то обозначением его класса точности является число из вышеприведенного ряда (1.8), выраженное в процентах.

**Пример:** если  $\gamma \pm 0,005$ , то класс точности будет обозначаться 0,5.

**Относительная основная погрешность**  $\delta$  для данного средства измерений нормируется, когда результат измерения  $X$  не может быть равен нулю. Используются два основных способа задания относительной основной погрешности.

Чаще всего это одночленная формула

$$\delta = \Delta/X = \pm q. \quad (1.9)$$

Здесь  $q$  есть число из того же ряда (1.8). В случае использования одночленной формулы (1.9) класс точности средства измерений записывается в виде величины  $q$ , выраженной в процентах и расположенной внутри окружности.

**Пример:** при  $d = 0,005$  запись класса точности выглядит как  $\textcircled{0,5}$ .

Иногда применяется двучленная формула

$$\delta = \frac{\Delta}{X} = \pm \left[ c + d \left( \left| \frac{X_k}{X} \right| - 1 \right) \right], \quad (1.10)$$

где  $X$  — измеряемая величина;  $X_k$  — конечное значение диапазона измерений;  $c$ ,  $d$  — числа из приведенного выше для  $p$  ряда (1.8). При использовании двучленной формулы (1.10) класс точности средства измерений записывается в виде дроби  $c/d$ , причем величины  $c$  и  $d$  выражаются в процентах.

**Пример:** если основная погрешность равна  $\delta = \frac{\Delta}{X} = \pm \left[ 0,004 + 0,0001 \left( \left| \frac{X_k}{X} \right| - 1 \right) \right]$ , то класс точности будет записан как 0,4/0,01.

Другие способы обозначения класса точности встречаются крайне редко.

#### 1.5.4. Проверка и калибровка средств измерений

**Проверка средства измерений** — совокупность операций, выполняемых аккредитованными в установленном порядке в области обеспечения единства измерений юридическими лицами и индивидуальными предпринимателями с целью определения и подтверждения соответствия средства измерений установленным техническим требованиям. Существует несколько способов проверки, различающихся для мер и измерительных приборов.

Для проверки мер используются следующие способы:

- сличение с более точной, чем поверяемая, образцовой мерой с помощью компарирующего прибора;
- измерение величины, воспроизводимой поверяемой мерой, измерительными приборами соответствующего разряда и класса («градуировка мер»).

Для наборов мер (или многозначных мер) часто используется *калибровка*, заключающаяся в сличении одной меры из набора (или одной из отметок шкалы многозначной меры) с более точной мерой. При этом размеры других мер поверяемого набора (или значения воспроизводимой величины на других отметках шкалы) определяют, сравнивая их в различных сочетаниях на приборах сравнения и обрабатывая полученные результаты.

*Измерительные приборы* могут быть поверены двумя способами:

- путем измерения с их помощью величины, воспроизводимой эталонами соответствующего разряда. Обычно при этом значения измеряемой величины выбирают равными соответствующим отметкам шкалы прибора, и основная погрешность равна наибольшей разности между результатом измерения и размером меры.

**Пример:** проверка весов взвешиванием образцовой гири (меры);

- измерением поверяемым и более точным (образцовым) прибором одной и той же величины (величин). При этом погрешность поверяемого прибора определяется разностью показаний поверяемого и образцового приборов.

**Пример:** проверка термометра с помощью образцового термометра путем измерения ими температуры воды в одном термостате.

Для сложных средств измерений, состоящих из нескольких более или менее самостоятельных измерительных блоков, возможна как комплексная поверка (при которой оно поверяется как целое), так и поэлементная поверка, при которой каждый их блоков поверяется отдельно.

Как сказано выше, при поверке важен выбор соотношения погрешностей образцового и поверяемого средств измерений. Его делают в соответствии с фундаментальным принципом пренебрежения малыми погрешностями (см. разд. 2.6.1). Обычно соотношение погрешностей выбирают в диапазоне  $(1:3 \div 1:5)$ , но иногда (с учетом конкретных особенностей процедуры поверки и требований к ней) используют и другие соотношения.

Осуществляют *первичную, периодическую, внеочередную и инспекционную* поверки.

*Первичная поверка* проводится всегда при выпуске средства измерений.

*Периодические поверки* осуществляются в установленные технической документацией сроки (чаще всего устанавливается *межповерочный интервал*) в течение всего времени эксплуатации и хранения средства измерений.

К *внеочередным поверкам* прибегают в особых ситуациях, когда возникают сомнения в качестве измерений, при повреждении поверочного клейма, при утере документации и т.д.

**Пример.** Внеочередная поверка весов после падения на них мойщика окон.

*Инспекционные поверки* проводятся при ревизиях.

Положительные результаты поверки удостоверяются *поверительным клеймом*, устанавливаемым на средстве измерений, и (или) *свидетельством о поверке*. В свидетельстве о поверке должен указываться эталон, к которому прослеживаются<sup>11</sup> результаты измерений, получаемые с помощью поверяемого средства измерений.

В специальных случаях проводятся также *ревизии и экспертизы* средств измерений.

Система поверок основана на следующих предположениях. Считается, что метрологические характеристики средств измерений:

- практически полностью определяют общую погрешность измерений;
- контролируются и подтверждаются в ходе периодических поверок;
- гарантированно сохраняются в течение межповерочного интервала.

На практике эти предположения часто не выполняются. Так, если при поверке выявлена недопустимая погрешность средства измерений, его регулируют либо изымают. Однако это означает, что какое-то время перед последней поверкой оно (средство измерений) имело неудовлетворительные метрологические характеристики. Кроме того, возможны кратковременные выходы процесса измерений из подконтрольного состояния.

Решением проблемы мог бы быть контроль стабильности средств измерений, аналогичный рассмотренному нами в гл. 4 (разд. 4.5.1 и сл.) контролю стабильности методов количественного химического анализа. При его применении можно было бы гарантировать статистическую подконтрольность работы средств измерений и

<sup>11</sup> Прослеживаемость — свойство результата измерений, заключающееся в возможности сопоставить этот результат с эталоном соответствующей величины через непрерывную цепь сличений (передачи размеров единиц величин). Подробнее о прослеживаемости см. разд. 3.5.1.

увеличить межповерочные интервалы (либо проводить поверки лишь при выходе средства измерений из подконтрольного состояния). Первые шаги в этом направлении уже делаются [15].

Об организационной стороне поверки см. 1.6.3.

**Калибровка средств измерений** — совокупность операций, выполняемых в целях определения действительных значений метрологических характеристик средств измерений [2]. Она выполняется в добровольном порядке преимущественно метрологическими службами юридических лиц с использованием эталонов, включенных в систему передачи размера (см. выше), причем заниматься этим могут юридические лица и индивидуальные предприниматели, как аккредитованные, так и не аккредитованные в области обеспечения единства измерений. В России деятельность по калибровке регламентирована законом [2] и рядом подзаконных актов. Однако в документах не всегда четко определено, какие именно средства измерений подлежат поверке, а какие калибровке, и в каких случаях.

## 1.6. Российская система измерений

В России в настоящее время эксплуатируются более миллиарда средств измерений (мер, измерительных приборов и т.д.), с помощью которых ежедневно проводится несколько миллиардов разнообразных измерений. Обеспечение надлежащего уровня точности измерений при таких масштабах требует соответствующего метрологического обеспечения на национальном уровне. А переход страны на инновационный путь развития на основе генерации знаний высокого уровня, их распространения и эффективного применения, ускорения нововведений, повышение экспортного потенциала отечественной продукции требуют также значительной модернизации существующих и создания новых государственных эталонов, измерительных средств для перспективных отраслей экономики, а также разработки комплекса нормативно-правовых и нормативно-методических документов.

Основным документом, определяющим государственную политику в области обеспечения единства измерений, является закон [2]. В нем определена сфера государственного регулирования в этой области, включающая измерения, выполняемые при:

- осуществлении деятельности в области здравоохранения и ветеринарной деятельности;
- осуществлении деятельности в области охраны окружающей среды и безопасности при чрезвычайных ситуациях;
- выполнении работ по обеспечению безопасных условий и охраны труда, а также при осуществлении производственного контроля за соблюдением установленных законодательством РФ требований промышленной безопасности к эксплуатации опасных производственных объектов;
- торговле и товарообменных операциях, расфасовке товаров;
- выполнении государственных учетных операций, проведении банковских, налоговых и таможенных операций;
- оказании услуг почтовой связи и учете объема оказанных услуг электросвязи операторами связи;

- осуществлении деятельности в области обороны и безопасности государства;
- осуществлении геодезической и картографической деятельности, в области гидрометеорологии;
- выполнении работ по оценке соответствия продукции, а также иных объектов установленным законами РФ обязательным требованиям;
- проведении официальных спортивных соревнований, обеспечении подготовки спортсменов высокого класса;
- выполнении поручений суда, органов прокуратуры, государственных органов исполнительной власти;
- осуществлении мероприятий государственного контроля (надзора).

Сфера государственного регулирования обеспечения единства измерений распространяется также на единицы величин, эталоны единиц величин, стандартные образцы и средства измерений, к которым установлены обязательные требования. К сфере государственного регулирования обеспечения единства измерений относятся также измерения, предусмотренные законом [16].

*Во всех остальных сферах положения закона [2] применяются в добровольном порядке.*

На рис. 1.2 показаны основные элементы структуры Российской системы измерений, функционирующей в настоящее время [17].

Государственная политика и нормативно-правовое регулирование в области обеспечения единства измерений лежит на Министерстве промышленности и торговли Российской Федерации (Минпромторг России), а федеральным органом исполнительной власти, осуществляющим функции по оказанию государственных услуг и управлению государственным имуществом в области обеспечения единства измерений — на Федеральном агентстве по техническому регулированию и метрологии (Росстандарт), который определяет и руководит работой своих территориальных органов.



Рис. 1.2. Основные элементы структуры Российской системы измерений

*Минпромторг РФ, Росстандарт и другие федеральные органы исполнительной власти*, осуществляющие государственные функции в области обеспечения единства измерений и государственного метрологического надзора осуществляют следующую деятельность в области обеспечения единства измерений:

- разрабатывают государственную политику и нормативно-правовое регулирование, координируют деятельность по нормативно-правовому регулированию;
- организуют взаимодействие с органами государственной власти иностранных государств и международными организациями;
- реализуют государственную политику и координируют деятельность по ее реализации в области обеспечения единства измерений;
- осуществляют государственный метрологический надзор и координируют деятельность по его осуществлению.

Полномочия между ними распределяет Правительство РФ.

Основными задачами *государственных научных метрологических институтов* являются:

- проведение фундаментальных и прикладных научных исследований, экспериментальных разработок и осуществление научно-технической деятельности в области обеспечения единства измерений;
- разработка, совершенствование, содержание, сличение и применение государственных первичных эталонов единиц величин, передача единиц величин от этих эталонов<sup>12</sup>;
- участие в разработке проектов нормативных документов в области обеспечения единства измерений, обязательная метрологическая экспертиза проектов нормативных правовых актов РФ, содержащих требования к измерениям, стандартным образцам и средствам измерений;
- создание и ведение Федерального информационного фонда по обеспечению единства измерений и предоставление содержащихся в нем документов и сведений;
- участие в международном сотрудничестве в области метрологии.

Нормативной базой решения этих задач является *Государственная система обеспечения единства измерений (ГСОЕИ)*, представляющая собой комплекс нормативно-технических документов, устанавливающих единую номенклатуру, способы представления и оценки метрологических характеристик средств измерений; требования к государственным испытаниям, поверке, ревизии и экспертизе средств измерений; правила стандартизации и аттестации методик измерений. Основой ГСОЕИ являются стандарты (ГОСТ и ГОСТ Р), а также правила (ПР) и рекомендации (Р) по метрологии<sup>13</sup>.

*Государственные региональные центры метрологии* проводят поверку средств измерений в соответствии с установленной областью аккредитации. Они совершенствуют, содержат и применяют государственные эталоны единиц величин, используемых для обеспечения прослеживаемости других эталонов единиц величин и средств

<sup>12</sup> Россия унаследовала практически всю эталонную базу СССР. В настоящее время она оценивается в несколько миллиардов долларов США и со временем только дорожает.

<sup>13</sup> После 1992 г. многие стандарты утверждаются Межгосударственным советом по стандартизации, метрологии и сертификации и действуют во всех государствах СНГ.

измерений к государственным первичным эталонам единиц величин. В целом тем же занимаются аккредитованные в установленном порядке в области обеспечения единства измерений юридические лица и индивидуальные предприниматели.

*Метрологические службы федеральных органов исполнительной власти* осуществляют метрологическую деятельность в рамках своих ведомств. Заметим, что не все предусмотренные законом [2] и показанные на рис. 1.2 метрологические службы федеральных органов существуют на самом деле.

О государственных метрологических службах (ГССД, ГССО и ГСВЧ) см. ниже, разд. 1.6.4.

Согласно Закону о техническом регулировании [16] все технические документы, за исключением регламентов, не носят обязательного характера. В то же время большинство метрологических требований выполняется либо в соответствии с законом [2], либо в добровольном порядке.

### *1.6.2. Утверждение типа средств измерений*

Средствам измерений на государственном уровне придается огромное значение. Действительно, от их состояния зависит успешность большинства отраслей народного хозяйства. Одним из видов Государственного регулирования в области измерений является утверждение типа средств измерений. Оно проводится в целях обеспечения единства измерений в стране.

*Утверждение типа средств измерений* — документально оформленное в установленном порядке решение о признании соответствия типа средств измерений метрологическим и техническим требованиям (характеристикам) на основании результатов их испытаний. Обязательному утверждению подлежат типы средств измерений, используемых в сфере государственного регулирования [2], хотя в добровольном порядке может утверждаться и тип средств измерений, используемый в других сферах. При утверждении типа средств измерений устанавливаются показатели точности, интервал между поверками средств измерений, а также методика поверки данного типа средств измерений.

Порядок проведения испытаний средств измерений в целях утверждения их типа, утверждения типа средств измерений и выдачи соответствующих свидетельств и т.д. устанавливаются соответствующими федеральными органами. На каждый экземпляр средств измерений утвержденного типа и на сопроводительные документы к указанным средствам измерений наносится знак утверждения их типа. Сведения об утвержденных типах средств измерений вносятся в Федеральный информационный фонд по обеспечению единства измерений.

Испытания стандартных образцов или средств измерений в целях утверждения типа проводятся юридическими лицами, аккредитованными в установленном порядке в области обеспечения единства измерений — государственными центрами испытаний средств измерений (ГЦИ СИ) — по заявке разработчика средства измерений.

В настоящее время порядок проведения испытаний и утверждения типа средств измерений включает [18]:

1. Рассмотрение ГЦИ СИ заявки с принятием решения о возможности проведения испытаний;



2. Испытания средства измерений в ГЦИ СИ, в ходе которых:
  - определяются метрологические и технические характеристики средства измерений;
  - идентифицируется программное обеспечение и оценивается его влияние на метрологические характеристики средства измерений;
  - разрабатывается или выбирается, а также опробуется методика поверки;
  - определяется интервал между поверками и т.д.
3. Разработку ГЦИ СИ описания типа средства измерений, утверждение (при необходимости) методики поверки, оформление акта испытаний средства измерений.
4. Утверждение (при положительных результатах испытаний) типа средств измерений Росстандартом с выдачей соответствующего свидетельства и внесением сведений о средстве измерений в информационную базу. Срок действия свидетельств для средств измерений серийного производства составляет 5 лет и может быть продлен без проведения испытаний еще на 5 лет решением Росстандарта.

В дальнейшем соответствие средств измерений утвержденному типу подлежит контролю, который проводится периодически в течение всего срока выпуска данного средства измерений.

Аналогичные процедуры применяются при утверждении типов стандартных образцов (см. разд. 3.2).

### *1.6.3. Метрологический надзор*

Государственный метрологический надзор осуществляется в сфере государственного регулирования обеспечения единства измерений (см. выше) в соответствии с законом [2]. Надзору подлежат:

- соблюдение обязательных требований в сфере государственного регулирования обеспечения единства измерений к измерениям, единицам величин, а также к эталонам единиц величин, стандартным образцам, средствам измерений при их выпуске из производства, ввозе на территорию Российской Федерации, продаже и применении на территории Российской Федерации;
- наличие и соблюдение аттестованных методик (методов) измерений;
- соблюдение обязательных требований к отклонениям количества фасованных товаров в упаковках от заявленного значения.

Государственный метрологический надзор распространяется на деятельность юридических лиц и индивидуальных предпринимателей, осуществляющих:

- измерения, относящиеся к сфере государственного регулирования обеспечения единства измерений;
- выпуск из производства предназначенных для применения в сфере государственного регулирования обеспечения единства измерений эталонов единиц величин, стандартных образцов и средств измерений, а также их ввоз на территорию Российской Федерации, продажу и применение на территории Российской Федерации;
- расфасовку товаров.



Государственный метрологический надзор и порядок его проведения устанавливается федеральными органами исполнительной власти и осуществляется должностными лицами федеральных органов исполнительной власти, осуществляющих государственный метрологический надзор, и их территориальных органов. В ходе надзора проверяется:

- соответствие используемых единиц величин единицам величин, допущенным к применению в Российской Федерации;
- состояние и применение эталонов единиц величин, стандартных образцов и средств измерений в целях установления их соответствия обязательным требованиям;
- наличие и соблюдение аттестованных методик (методов) измерений;
- соблюдение обязательных требований к измерениям и обязательных требований к отклонениям количества фасованных товаров в упаковках от заявленного значения.

При выявлении нарушений должностное лицо, осуществляющее государственный метрологический надзор, обязано принимать предусмотренные законодательством меры по устранению этих нарушений.

#### ***1.6.4. Государственные метрологические службы [2]***

Существуют три государственные метрологические службы.

***Государственная служба времени, частоты и определения параметров вращения Земли.*** В сферу ее деятельности входит выполнение следующих основных задач:

- воспроизведение размеров единиц времени и частоты;
- определение параметров вращения Земли;
- формирование и хранение национальной шкалы времени;
- передача сигналов времени и частоты;
- определение соотношения между атомной шкалой времени и шкалой всемирного времени;
- обеспечение международного единства измерений времени и частоты.

***Государственная служба стандартных образцов состава и свойств веществ и материалов*** осуществляет деятельность по разработке, испытанию и внедрению стандартных образцов состава и свойств веществ и материалов в целях обеспечения единства измерений на основе применения указанных стандартных образцов, а также по ведению соответствующих разделов Федерального информационного фонда по обеспечению единства измерений. Подробнее о стандартных образцах состава см. разд. 3.2.

***Государственная служба стандартных справочных данных о физических константах и свойствах веществ и материалов*** осуществляет деятельность по разработке и внедрению стандартных справочных данных о физических константах и свойствах веществ и материалов в науке и технике в целях обеспечения единства измерений на основе применения указанных стандартных справочных данных, а также по ведению соответствующих разделов Федерального информационного фонда по обеспечению единства измерений.

***Государственная служба стандартных образцов состава и свойств веществ и материалов*** осуществляет деятельность по разработке, испытанию и внедрению

стандартных образцов состава и свойств веществ и материалов в целях обеспечения единства измерений на основе применения указанных стандартных образцов, а также по ведению соответствующих разделов Федерального информационного фонда по обеспечению единства измерений.

### *1.6.5. Международные метрологические организации*

В настоящее время большинство стран тесно связаны между собой — развитая торговля, кооперация в сфере производства требуют обеспечения сопоставимости результатов измерений не только в пределах одного государства, но и во всем мире. Принцип единства измерений и требований к методам измерений, к техническим характеристикам измерительных приборов, к обеспечению средств измерений методами и средствами поверки, их испытаний и контроля качества составляет основу единой метрологической системы, устраняющей технические барьеры в международном научно-техническом и промышленном сотрудничестве. Единый метрологический подход способствует достижению максимальной экономии средств, затрачиваемых на научные исследования и создание материальной базы для метрологической деятельности.

Для решения задачи обеспечения сопоставимости результатов измерений на Земле созданы *международные метрологические организации*. Благодаря их усилиям в большинстве стран мира принята Международная система СИ, действует сопоставимая терминология, приняты рекомендации по способам нормирования характеристик средств измерения, по сертификации и испытаниям средств измерений и т.д.

Старейшей международной метрологической организацией является Международная организация мер и весов (МОМВ), основанная в 1875 г. с целью унификации применяемых в разных странах систем единиц измерения, установления единообразия эталонов длины и массы. В настоящее время МОМВ, кроме единиц длины и массы, занимается системами единиц времени и частоты, а также электрическими, фотометрическими, стабилизированными лазерными, гравитационными, термометрическими и радиометрическими измерениями. В состав МОМВ входят *Генеральная конференция по мерам и весам (ГКМВ)*, *Международное бюро мер и весов (МБМВ)*, *Международный комитет по мерам и весам (МКМВ)*. Для химической метрологии особенно важна работа *Консультативного комитета по количеству вещества (Метрология в химии)* при МКМВ.

С целью выработки единой стратегии в области метрологии и проведения ее в жизнь в 1955 г. была создана *Международная организация законодательной метрологии (МОЗМ)*. В настоящее время в ее состав входят 114 государств (58 государств-членов и 56 государств-членов-корреспондентов), и это число постоянно растет. Ее задачами и функциями являются:

- установление взаимного доверия к результатам измерений технических характеристик различных материалов каждой из стран-участниц;
- выбор необходимых и достаточных характеристик и требований, которым должны соответствовать средства измерений, чтобы их можно было использовать во всех странах;

- оказание помощи нуждающимся странам в организации национальных метрологических служб и их техническом оснащении, в подготовке и повышении квалификации специалистов-метрологов;
- создание систем сертификации и управления качеством, а также проведение исследований методов поверки, испытаний и контроля средств измерений;
- организация сличений на уровне рабочих эталонов и средств измерений, разработка новых методов метрологического надзора для контроля в процессе сертификации у изготовителя средств измерений, в том числе на основе методологии управления качеством, в соответствии со стандартами ИСО (см. ниже) серии 9000.

Высший орган МОЗМ представлен *Международной конференцией законодательной метрологии*, рабочим органом которой является *Международный комитет законодательной метрологии (МКЗМ)*, а исполнительным органом — *Международное бюро законодательной метрологии (МБЗМ)*. МБЗМ является Центром документации по законодательной метрологии. Разработка проектов рекомендаций и документов осуществляется техническими комитетами и подкомитетами, которые делятся на три основные группы (по общим вопросам метрологии, по отдельным видам измерений, по метрологическому обеспечению).

Функцию национального Секретариата МОЗМ России выполняет ВНИИ метрологической службы (ВНИИМС).

МОЗМ тесно связан с МБМВ, *Международной организацией стандартизации (ИСО)* и *Международной электротехнической комиссией (МЭК)* (см. ниже). Цель этого сотрудничества — избежать дублирования, упущений и расхождений при разработке технической документации.

В 1988 г. подписана Конвенция об образовании *Общеввропейской метрологической организации (EUROMET, ЕВРОМЕТ)*. Основные функции ее деятельности составляют исследование и разработка эталонов единиц измерений; исследования по разработке первичных эталонов; развитие поверочных служб на высшем метрологическом уровне и методов наивысшей точности.

Региональная метрологическая организация *КООМЕТ* (кооперация в метрологии) была создана в 1991 году. В основном в нее входят страны бывшего СЭВ и страны, образовавшиеся после распада СССР. В 2000 году организация была официально переименована в Евро-Азиатское сотрудничество государственных метрологических учреждений. Цели КООМЕТ: содействие эффективному решению вопросов единообразия мер, единства измерений и требуемой точности их результатов; содействие развитию сотрудничества национальных экономик и устранению технических барьеров в международной торговле; сближение деятельности метрологических служб евро-азиатских стран с деятельностью аналогичных служб других регионов. Решения КООМЕТ носят рекомендательный характер.

Для координации метрологической деятельности в рамках СНГ создана *Научно-техническая комиссия по метрологии (НТКМетр)* — постоянно действующий рабочий орган Межгосударственного совета по стандартизации, метрологии и сертификации. Основная его задача — выработка предложений по формированию согласованной научно-технической политики и координации работ по метрологии

в государствах-участниках «Соглашения о проведении согласованной политики в области стандартизации, метрологии и сертификации», подписанного главами правительств государств-участников СНГ 13 марта 1992 в г. Москве.

Из метрологических организаций следует отметить также *Международную конфедерацию по измерительной технике (ИМЕКО)*, в составе которой имеется технический комитет по метрологии (ТК-8).

Специальными вопросами метрологии и измерительной техники занимается также ряд других международных организаций: *Консультативный комитет по радиосвязи (МККР)*, *Консультативный комитет по телефонии и телеграфии (МККТТ)*, *Международная организация гражданской авиации (ИКАО)*, *Агентство по атомной энергии (МАГАТЭ)*, *Комитет по исследованию космического пространства (КОСПАР)*.

Международные метрологические организации активно взаимодействуют с организациями, действующими в смежных областях, из которых выделим Международную организацию по стандартизации и Международную энергетическую комиссию.

*Международная организация по стандартизации* была основана в 1946 г. в Лондоне двадцатью пятью национальными организациями по стандартизации. При выборе ее названия использовали греческое слово *isos* — равный, поэтому аббревиатура наименования звучит одинаково на всех языках: *ISO*, по-русски *ИСО*. Сфера деятельности ИСО касается стандартизации во всех областях, кроме электротехники и электроники, относящихся к компетенции Международной электротехнической комиссии (МЭК, см. ниже), хотя ряд работ выполняется этими организациями совместно. Кроме стандартизации ИСО занимается проблемами сертификации.

ИСО определяет свои задачи как содействие развитию стандартизации и смежных видов деятельности в мире с целью обеспечения международного обмена товарами и услугами, а также развитие сотрудничества в интеллектуальной, научно-технической и экономической областях. В состав ИСО в настоящее время входит более 120 государств — все промышленно развитые страны и многие развивающиеся. Члены ИСО делятся на две категории: активные участники и наблюдатели. В качестве последних выступают члены, уплачивающие ограниченные взносы в бюджет ИСО. Они имеют право на получение всех издаваемых международных стандартов (МС), а также право участвовать в различных заседаниях, не имея, однако, решающего голоса. Активные члены обязаны принимать постоянное участие в разработке рабочих документов и выражать свое мнение путем голосования.

В структуре ИСО около 160 технических комитетов, 600 подкомитетов и 300 рабочих групп, которые разрабатывают стандарты практически во всех сферах человеческой деятельности. Каждая страна может быть представлена одной организацией. От СССР членом ИСО с 1949 г. являлся Государственный комитет СССР по стандартизации. В настоящее время от Российской Федерации членом ИСО является Федеральное агентство по техническому регулированию и метрологии.

Международные стандарты ИСО, как и других международных организаций, формально являются добровольными к применению. Как правило, они аккумуля-

лируют в себе передовой опыт стандартизации и влияют на торговые, культурные обмены между странами. Поэтому практически все государства стремятся в максимально высокой степени и оперативно вводить стандарты ИСО целиком или частично в национальные нормативно-технические документы.

*Международная электротехническая комиссия (IEC, МЭК)* — международная некоммерческая организация по стандартизации в области электрических, электронных и смежных технологий. МЭК была основана в 1906 году и в настоящее время в ее состав входят более 70 стран. МЭК составлена из представителей национальных служб стандартов. МЭК способствовала развитию и распространению стандартов для единиц измерения, особенно гаусса, герца и вебера. Именно МЭК предложила систему стандартов, которая в конечном счете стала единицами СИ. Издаваемый ею Международный электротехнический словарь — тезаурус для электрических и электронных отраслях промышленности.

Стандарты МЭК имеют номера в диапазоне 60000-79999, и их названия имеют вид *МЭК [номер стандарта][название стандарта]*, например IEC 60584-1 (1995-09) Thermocouples — Part 1: Reference tables. Стандарты, разработанные МЭК совместно с ИСО, имеют названия вида *ISO/IEC [номер стандарта][название стандарта]*, например ISO/IEC 17025:2005 «General requirements for the competence of testing and calibration laboratories».

Метрологией химического анализа занимаются международные организации, специализирующиеся на аналитической химии, из которых выделим *IUPAC* и *EURACHEM*.

Международный союз теоретической и прикладной химии (*ИЮПАК, IUPAC*) — основанная в 1919 г. международная организация, членами которой являются научные учреждения более чем 40 стран. Помимо самых общих вопросов (терминологии, символики, единиц измерения) ИЮПАК занимается созданием эталонов и стандартных методов для научных исследований.

*EURACHEM (ЕВРАХИМ)* — основанная в 1989 г. сеть организаций в Европе, имеющих целью создание системы международного контроля измерений и повышение качества измерений в аналитической химии. *ЕВРАХИМ* создает рабочие группы по наиболее актуальным проблемам в этой области, выпускает справочники, организует технические семинары, координирует практическую работу в области аналитических измерений.

Активно работают международные организации, занимающиеся сертификацией продукции и аккредитацией лабораторий. Это прежде всего *ISO* — *международная организация по сертификации*; *ILAC* — *международное сотрудничество по аккредитации лабораторий и контрольных (инспекционных) органов*. Упомянем также *Европейскую организацию по испытаниям и сертификации (ЕОИС)* — метрология играет важнейшую роль в процедурах сертификации продукции и аккредитации лабораторий.

В работе большинства этих организаций наша страна принимает активное участие. В частности, Россия участвует во многих двух- и многосторонних договоренностях о взаимном признании результатов испытаний, утверждении типов и поверок.

## 1.7. Количественный химический анализ и особенности его метрологии

### 1.7.1. Что измеряют при КХА

Существуют два подхода к определению того, что именно измеряют при КХА.

**Вариант 1.** Говорят, что при химическом анализе измеряют *концентрацию* или *содержание* тех или иных веществ.

*Концентрация — величина, характеризующая относительное содержание данного компонента в многокомпонентной термодинамической системе* [19].

При этом концентрация может быть выражена в разных единицах, как в производных единицах системы СИ (моль/м<sup>3</sup>, кг/м<sup>3</sup>, концентрация числа частиц 1/м<sup>3</sup> и т.д.), так и с использованием внесистемных единиц (например, в долях). В этом случае используются формулировки типа «измерение концентрации бензола в воде» с возможным уточнением «выраженной в объемных процентах». Часто, чтобы избежать терминологического конфликта, вместо термина «измерение» применяют близкий по смыслу термин «определение» («определение концентрации мочевины в воде»).

Преимущества такого подхода — возможность использования любых единиц, часто специфических для конкретных областей (так, в учении о растворах пользуются понятием моляльность (число молей вещества, растворенного в 1000 г растворителя)). Кроме того, с позиций обучения наличие единого понятия «концентрация» естественно подчеркивает, что речи идет в целом об одном и том же. Недостаток — понятие «концентрация» выбивается из ряда других величин; приходится признавать особую ситуацию в аналитической химии по сравнению с другими измерениями.

**Вариант 2.** Выделяют несколько разных величин (чаще всего это массовая или объемная концентрация, концентрация количества вещества и концентрация числа частиц) и считают, что эти величины измеряются при КХА. Тогда можно говорить об «измерении объемной концентрации (доли) бензола в бензине».

Достоинством этого способа является его близость к описанию физических величин. Иногда говорят, что есть разные величины, а концентрация — обобщающее их название, аналогично тому, что вязкость — обобщающее название нескольких величин, таких как кинематическая и динамическая вязкость. Однако в аналитической химии ситуация иная — если кинематическая и динамическая вязкость измеряются по-разному, то концентрация измеряется одинаково, как ее ни называй.

Использование такого подхода имеет важный недостаток, особенно при обучении. Создается иллюзия, что в зависимости от того, в каких единицах выражается результат измерения, измерения проводятся по разным методикам. Между тем с точки зрения химика при выполнении одних и тех же манипуляций (отбор пробы, ее растворение, концентрирование, измерение аналитического сигнала и т.д.) выполняется одна и та же методика, вне зависимости от того, в каких единицах выражается результат. Если необходимо выразить результат КХА в других единицах, это легко сделать путем пересчета.

Кроме того, возникает необходимость говорить о множестве несистемных величин, что еще более затрудняет изложение вопроса.

*Содержание* – общий термин, включающий как концентрацию, так и количество вещества.

*Пример.* Массу железа в железнодорожном вагоне (в пересчете на оксид) можно назвать «содержанием железа в вагоне».

Мы будем использовать термин «измерение» в общем смысле, а в примерах говорить об определении концентрации тех или иных веществ.

### *1.7.2. Общий взгляд на метрологию химического анализа*

Согласно определению, данному в разд. 1.1.1, химическая метрология занимается измерениями в количественном химическом анализе. Метрология других измерений, проводимых в химии (измерения массы, давления и т.д.), не имеет существенных особенностей по сравнению с аналогичными измерениями в других областях.

Количественный химический анализ, целью которого является измерение концентрации различных веществ в объектах, представляет собой весьма специфическую область измерений. Длительное время велась (да и сейчас ведется) дискуссия о том, является ли количественный химический анализ измерением. К настоящему времени мнение большинства специалистов сводится к тому, что эта область химии является измерением, однако измерением особого типа, метрология которого требует специфических подходов. С учетом этого существуют особые стандарты и другие документы, регламентирующих метрологические аспекты количественного химического анализа. Аналогичная деятельность ведется и в международном масштабе.

Конечной стадией количественного химического анализа почти всегда является прямое измерение физической величины: массы в гравиметрическом методе, оптических величин в спектральных методах, электрических величин в электрохимических методах и т.д. Однако это измерение крайне редко является определяющим с точки зрения общей погрешности результата анализа. В наибольшей степени общая погрешность определяется другими стадиями анализа – процедурами отбора пробы, пробоподготовки, перевода определяемого вещества в аналитическую форму и т.д., то есть теми стадиями, которые и делают измерение собственно анализом. С учетом этого обстоятельства результат анализа чаще всего рассчитывают с использованием эмпирически определяемой зависимости, связывающей аналитический сигнал – результат измерения той или иной физической величины (величин) или функцию таких результатов – с концентрацией или количеством определяемого компонента в образцах сравнения.

*Пример.* Концентрацию вещества в растворе, определяемую методом спектрофотометрии, находят либо на основании уравнения  $C = A/\varepsilon l$  (где  $\varepsilon$  – молярный коэффициент поглощения, л·моль<sup>-1</sup>·см<sup>-1</sup>;  $l$  – толщина поглощающего слоя) или с помощью градуировочной зависимости оптической плотности  $A$  от концентрации  $C$  (детальнее см. разд. 3.1 и сл.).

Из сказанного ясно, что количественный химический анализ почти всегда представляет собой косвенное измерение<sup>14</sup>.

Другой особенностью количественного химического анализа как процесса измерения является большая роль человеческого фактора, поскольку результаты многих важных этапов химического анализа (отбор проб, дозирование реактивов, перемешивание смесей веществ и т.д.) прямо зависят от квалификации, степени ответственности и даже настроения выполняющего эти процедуры сотрудника. Для удовлетворительного выполнения многих видов химического анализа нужны лаборанты, не только прошедшие соответствующее обучение, но и имеющие достаточный опыт подобной работы. Один из путей решения проблемы точного и надежного химического анализа — автоматизация всех стадий анализа. Однако в обозримом будущем доля ручного труда останется высокой — многие операции невозможно или слишком дорого автоматизировать. Кроме того, даже полностью автоматизированные сложные системы подвержены сбоям разного рода, зачастую скрытым от оператора. Именно поэтому столь большое значение придается как аттестации методик КХА и оценке пригодности ее в конкретной лаборатории (валидации), так и последующему внутрилабораторному контролю и внешней оценке качества результатов анализа, аккредитации химических лабораторий, а при обработке результатов количественного химического анализа — грубым погрешностям измерений.

И, наконец, определяющим метрологию количественного химического анализа фактором является наличие в огромном большинстве методов химического анализа объектов сложного состава так называемых *скрытых, или не выявленных систематических погрешностей*. Это разновидность систематических погрешностей, которые по тем или иным причинам не могут быть выявлены и скорректированы в ходе или по окончании измерений. Все это заставляет искать особые пути построения метрологических схем для количественного химического анализа.

При постановке проблемы построения для количественного химического анализа схемы передачи размера, подобной показанной на рис. 1.1, необходимо ответить на следующие вопросы:

- что может быть первичным эталоном в анализе вещества?
- что представляют собой эталоны в химическом анализе?
- что выполняет роль средства измерений в химическом анализе?

В отношении первого вопроса приведем следующие соображения. За единицу количества вещества в химии принят моль, с 2019 г. определяемый как количество вещества, содержащее число структурных элементов этого вещества, равное числу Авогадро. Из этого определения ясно, что создать единый для разных веществ эталон моля непросто: в определении задано лишь число структурных элементов (при количественном химическом анализе — обычно атомов или молекул) в одном моле. Поэтому в качестве эталонов в химическом анализе используется множество *стандартных образцов*, которые могут представлять собой как чистые вещества, так и их растворы, а также материалы сложного состава, в которых тем или иным способом определена концентрация определяемых веществ (*аналитов*).

<sup>14</sup> См. также примечание 1 к Введению.



Необходимость использования множества стандартных образцов сложного состава связана с тем, что для исследования разных объектов используются разные методы. Кроме того, при использовании многих методик возникают *скрытые, или не выявленные, систематические погрешности*, связанные со сложностью решаемой задачи — измерением концентрации аналита, входящего в состав объекта сложного состава.

С учетом сказанного метрология химического анализа строится обычно не по классическим канонам, описанным выше, а специфическим образом. Как — описано в гл. 3.

## ГЛАВА 2

# ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Как показано в гл. 1, химический анализ представляет собой специфическую разновидность измерений. Согласно третьей аксиоме метрологии (см. разд. 1.1) результат измерения есть случайная величина. Поэтому к результатам измерений вообще и к результатам химического анализа в частности можно применять хорошо разработанный аппарат математической статистики. Действительно, методы математической статистики лежат в основе большинства расчетов, проводимых при обработке результатов измерений.

**Математическая статистика** — наука о математических методах систематизации и использования статистических данных для получения научных и практических выводов.

При этом под статистическими данными понимают сведения о тех или иных показателях совокупности объектов, т. е. описание массовых явлений.

**Примеры** статистических данных: результаты нескольких измерений концентрации перманганат-иона студентом в практикуме по аналитической химии; результаты измерений напряжения в сети разными вольтметрами; количественные факторы, связанные со здоровьем людей (рост, вес, артериальное давление и т. д.), измеренные у жителей какого-либо города; курсы акций на бирже за длительный период.

Математическая статистика представляет собой ветвь теории вероятностей.

Вопросами применения математических методов в химии занимается наука, называемая *хеометрикой*.

Огромная доля вычислений в химии представляет собой те или иные виды обработки результатов количественного химического анализа. Поэтому можно утверждать, что хеометрика в значительной степени занимается применением математической статистики к результатам, получаемым при анализе вещества. Здесь хеометрика смыкается с метрологией химического анализа.

Математическая статистика — старая наука, в арсенале которой имеется множество методов «на разные случаи жизни», из которых в хеометрике и химической метрологии используется лишь часть.

В данной главе приведены некоторые сведения из математической статистики, необходимые для дальнейшего рассмотрения метрологии и методов обеспечения качества химического анализа. Поэтому мы будем полагать, что обрабатываемый числовой материал представляет собой результаты измерений.

Предполагается, что читатель знаком с элементарными понятиями теории вероятностей.

## 2.1. Общие понятия математической статистики

### 2.1.1. Генеральная совокупность и выборка

При измерениях мы всегда имеем дело с некоторой совокупностью результатов измерений. Для всех видов статистических измерений, кроме совместных, предметом статистической обработки в простейшем случае являются результаты нескольких измерений одной и той же величины  $X$ , полученные в одних условиях («параллельные измерения»):

$$X_1, X_2, \dots, X_p, \dots, X_N. \quad (2.1)$$

В математической статистике эту совокупность называют *выборкой*, явно или неявно полагая, что существует *генеральная совокупность*, из которой эта выборка получена. Это — *основной постулат математической статистики*<sup>15</sup>.

*Генеральная совокупность*, соответствующая выборке (2.1), представляет собой воображаемую совокупность, состоящую из бесконечно большого числа результатов измерений  $X$ , описываемых *функцией распределения*<sup>16</sup>:

$$\Phi(X) \equiv P(\xi < X), \quad (2.2)$$

где  $\xi$  — случайная величина, соответствующая генеральной совокупности,  $P$  — вероятность. Пример функции распределения показан на рис. 2.1. Значение функции распределения в какой-либо точке  $X_i$ , таким образом, равно вероятности  $P_i$  того, что случайная величина принимает значение, меньшее  $X_i$ .

Очень важна также *функция плотности распределения*:

$$\varphi(X) = \frac{d(\Phi(X))}{dX}. \quad (2.3)$$

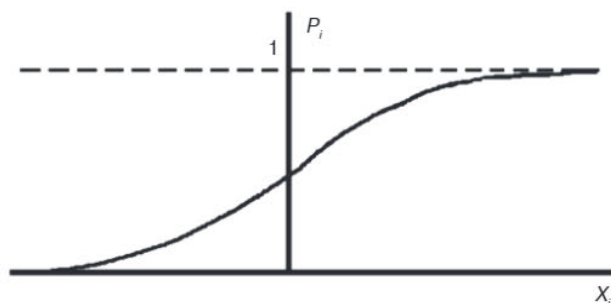


Рис. 2.1. Пример функции распределения. Вероятность того, что случайная величина  $\xi$  меньше  $X_i$ , равна  $P_i$

<sup>15</sup> Бывают и более сложные ситуации, когда и выборка, и соответствующая ей генеральная совокупность имеют сложную структуру. Приведенные здесь общие рассуждения в целом справедливы и для этих ситуаций.

<sup>16</sup> В отличие от большинства курсов математической статистики мы не рассматриваем редко встречающуюся при измерениях ситуацию, когда генеральная совокупность представляет собой конечное множество (например, разноцветные шары в «черном» ящике).

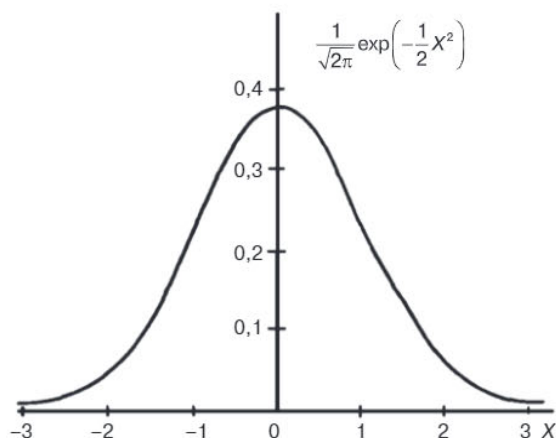


Рис. 2.2. Функция плотности нормального распределения при  $\mu = 0$  и  $\sigma = 1$

График функции плотности распределения (для нормального распределения, см. разд. 2.3.1.1) показан на рис. 2.2. Смысл его в том, что вероятность попадания значения случайной величины в какой-либо интервал равна площади (криволинейной) трапеции над этим интервалом. Полная площадь под кривой функции плотности распределения равна 1.

В рамках основного постулата математической статистики выборка как будто бы получается путем последовательного извлечения результатов из генеральной совокупности случайным образом. Можно сказать также, что генеральная совокупность, соответствующая выборке (т. е. результатам параллельных измерений), получилась бы, если бы мы выполнили бесконечное число измерений той же величины в тех же неизменных условиях. Другими словами, выборка бесконечно приближается к генеральной совокупности с ростом числа измерений при неизменности условий, в которых эти измерения выполняются.

Ниже вместо «генеральная совокупность» часто говорится просто «совокупность», а вместо «генеральная совокупность, имеющая нормальное распределение» — «нормальная совокупность».

Вообще математическая статистика — наука о том, как получать сведения о свойствах генеральных совокупностей по выборкам из них. Конечно, при использовании математической статистики в метрологии вообще и в метрологии химического анализа в частности предполагается, что свойства совокупности описывают свойства объекта и метода анализа. Так, в отсутствие систематических погрешностей математическое ожидание генеральной совокупности, соответствующей выборке (2.1) есть истинное значение измеряемой величины.

### 2.1.2. Ограниченность возможностей математической статистики. Статистические модели

Как сказано выше, основной задачей математической статистики является получение информации (оценка параметров, проверка гипотез) относительно свойств совокупности (совокупностей) по имеющейся выборке (выборкам). Однако вопрос



о связи параметров совокупности (и, соответственно, результатов обработки данных измерений методами математической статистики) с измеряемыми величинами сама математическая статистика не решает и решить не может в принципе. Действительно, при обработке результатов измерений предполагается, что они определяются собственно измеряемой величиной, а также определенными систематическими и случайными погрешностями измерений. Полная проверка этих предположений только посредством манипуляций с результатами измерений (выборкой) невозможна.

Поясним это на примере.

**Пример.** Измеряется концентрация стронция в магниевом сплаве, для чего выбран атомно-абсорбционный метод с пламенной атомизацией после растворения пробы в смеси минеральных кислот. Проведя несколько измерений в одних условиях, мы получим выборку, состоящую из результатов анализа, обрабатывая которую, можно рассчитать искомую концентрацию. Методы математической статистики позволяют исследовать форму распределения результатов анализа и найти параметры выборки (такие как среднее значение и стандартное отклонение), рассчитать доверительный интервал для найденного среднего значения. Однако вопрос о связи найденного среднего значения с истинной концентрацией стронция в сплаве при этом решен не будет. Мы обязаны предположить (на основе теоретических знаний, предшествующего опыта, специальных экспериментов и т. д.), что, во-первых, концентрация стронция в пробе и результаты атомно-абсорбционного анализа действительно связаны между собой и, во-вторых, процедура анализа не вносит существенных систематических погрешностей. В нашем примере последнее, возможно, и не выполняется — кальций, обычно содержащийся в магниевых сплавах, влияет на результаты определения стронция.

Проверка этих предположений только статистической обработкой результатов измерений концентрации стронция в сплаве без привлечения априорной информации, в частности, знания физических законов, и без проведения специально поставленных опытов невозможна (что вполне соответствует первой аксиоме метрологии!).

Простым и наглядным способом выражения исходных предположений об экспериментальных данных являются *статистические модели*.

При записи таких моделей в левой части указываются результаты измерений. При этом зачастую результаты сопровождаются индексами, соответствующими плану эксперимента.

**Пример.** Если модель описывает результаты нескольких взвешиваний одной и той же гири на нескольких весах и на каждом весе взвешивают только один раз, эти результаты можно обозначить как  $X_{ij}$ , где  $i$  — индекс прибора,  $j$  — номер измерения на  $i$ -м приборе).

В правой части модели находятся так называемые *эффекты* — формализованное выражение наших знаний и предположений о факторах, определяющих результаты измерений. В приведенном примере, скорее всего, результаты измерений определяются различиями между весами («эффект прибора») и случайной погрешностью измерений.

Систематическим погрешностям соответствуют систематические эффекты, а случайным погрешностям — случайные. Одни и те же экспериментальные данные могут обрабатываться в рамках нескольких моделей. Выбор конкретной модели определяется:

- целями исследователя (на какие именно вопросы он хочет получить ответы по результатам проведенных измерений);
- планом эксперимента;
- исходными предположениями о погрешностях измерений и их распределениях.

Простейшая выборка (2.1) в предположении о нормальном распределении результатов (см. разд. 2.3.1.1) и в отсутствие систематических погрешностей будет обрабатываться в рамках модели

$$X_i = \mu + \varepsilon, \quad (2.4)$$

где  $X_i$  – результат  $i$ -го измерения;  $\mu$  – постоянный эффект, равный математическому ожиданию совокупности и *соответствующий в отсутствие систематических погрешностей истинному значению измеряемой величины*;  $\varepsilon$  – нормально распределенная случайная величина  $\varepsilon \in N(0, \sigma_\varepsilon^2)$ .

Примеры более сложных моделей приведены ниже.

Записывая модель, мы *всегда* делаем предположения относительно распределений случайных погрешностей, их зависимости от измеряемых величин, вида эффектов и т. д. Справедливость модели может быть проверена либо непосредственно в ходе обработки результатов измерений, либо в специально поставленных экспериментах. Некоторые способы такой проверки рассмотрены ниже. Модель может изменяться (например, упрощаться путем исключения эффектов) по результатам проверки гипотез; примеры таких упрощений см. в гл. 4.

### 2.1.3. Оценка параметров. Некоторые свойства оценок

После того как выбор модели сделан, приступают собственно к обработке результатов измерений, которая состоит в получении оценок и проверке гипотез. О проверке гипотез говорится в разд. 2.1.4 и сл.; здесь речь пойдет об оценках.

Оценивают параметры, характеризующие генеральную совокупность, а именно: параметры распределений случайных величин, коэффициенты, описывающие силу связи между случайными величинами и т. д. Другими словами, по результатам измерений (выборке) оценивают параметры выбранной модели. Модель же, в свою очередь, выражает наши представления об «истине», т. е. о свойствах генеральных совокупностей.

Реально мы имеем дело с выборками, в силу их конечности отражающими свойства этих совокупностей лишь в той или иной степени. Поэтому мы можем находить лишь *оценки* параметров, т. е. рассчитывать параметры, или характеристики выборки (выборок), более или менее точно отражающие соответствующие параметры генеральной совокупности (совокупностей). Точные же значения параметров генеральной совокупности остаются неизвестными – высокопарно выражаясь, это связано с вероятностной природой нашего мира.

Существует развитая классификация оценок по их свойствам.

*Состоятельной* оценкой  $\hat{Y}(X_1, X_2, \dots, X_N)$  соответствующего параметра совокупности  $\eta$  называется такая оценка, которая сходится к  $\eta$  по вероятности при увеличении объема выборки  $N$ , т. е. если вероятность любого конечного отклонения  $|\hat{Y} - \eta|$  стремится к нулю при  $N \rightarrow \infty$ . Другими словами, состоятельная оценка стремится к соответствующему параметру совокупности при увеличении числа измерений. Все рассматриваемые ниже оценки являются состоятельными.



**Смещением** оценки  $\hat{Y}$  параметра  $\eta$  называется разность  $(MY - \eta)$ , где  $MY$  — математическое ожидание величины  $Y$ . Величина называется **несмещенной** оценкой параметра  $\eta$ , если  $(MY - \eta) = 0$ .

**Примеры.** Среднее арифметическое значение выборки есть несмещенная оценка математического ожидания нормальной совокупности. Стандартное отклонение выборки как оценка параметра  $\sigma$  нормальной совокупности — смещенная оценка (см. разд. 2.2).

В математической статистике рассматриваются и другие свойства оценок.

#### 2.1.4. Статистические гипотезы

**Гипотеза** есть исходное предположение о свойствах исследуемого объекта, проверяемое в ходе обработки результатов измерений методами математической статистики.

Подчеркнем, что гипотезы всегда формулируются относительно свойств исследуемого объекта, а не относительно свойств результатов измерений. В терминах математической статистики это означает, что гипотезы есть предположения относительно тех или иных свойств генеральных совокупностей (предполагается, что параметры этих совокупностей отражают соответствующие свойства реальных объектов, включая особенности процесса измерений). Гипотезы обозначают буквой  $H$  с тем или иным индексом; основную («нулевую») гипотезу обычно обозначают как  $H_0$ .

**Приведем пример.** В [20] в главе «статистические методы проверки гипотез» решается задача, названная (вероятно, для краткости) «сравнение двух средних». Экспериментальные данные представляют собой две независимые выборки вида (2.1), а средние значения, о которых идет речь, — средние элементов этих выборок. Однако очевидно, что сравнение средних значений не требует статистики — это просто два числа, из которых одно почти всегда больше другого. Гипотеза же проверяется относительно равенства неизвестных величин — математических ожиданий двух совокупностей, каждая из которых соответствует своей выборке. Поэтому мы, решая эту задачу, называем ее «проверка гипотезы о равенстве двух математических ожиданий» (см. разд. 2.5).

Реальные результаты измерений, являющиеся, в соответствии с основным постулатом математической статистики, выборками из генеральных совокупностей,отягчены случайными погрешностями и позволяют судить о свойствах исследуемого объекта лишь предположительно, с определенной вероятностью. Соответственно, при проверке любой гипотезы (при принятии решения, отвергается ли данная гипотеза или не отвергается) возможна ошибка. Различают ошибки первого и второго рода, которые при проверке некоторой гипотезы  $H_0$  можно представить следующим образом ([21]; цит. по [22]):

В действительности:	Если основная гипотеза:	
	не отвергается	отвергается
гипотеза $H_0$ верна	Правильное решение	Ошибка первого рода
гипотеза $H_0$ неверна	Ошибка второго рода	Правильное решение

Как видно, ошибку первого рода допускают, отвергая гипотезу, которая на самом деле верна; ошибку второго рода допускают, не отвергая неверную гипотезу<sup>17</sup>.

<sup>17</sup> Часто говорят также об «ошибках третьего рода», под которыми понимают ошибки при проверке гипотез, обусловленные неправильными предположениями о форме распределений.

При проверке гипотезы необходимо **заранее** определить, с какой вероятностью  $\alpha$  мы можем допустить ошибку первого рода и с какой вероятностью  $\beta$  — ошибку второго рода (в англоязычной литературе и компьютерных программах для вероятности ошибки первого рода используют также обозначение  $P$  от англ. *Probability*). Вероятности  $\alpha$  и  $\beta$  не являются независимыми — при уменьшении вероятности ошибки первого рода  $\alpha$  растет вероятность ошибки второго рода  $\beta$ . Заметим, что в большинстве случаев вероятность ошибки второго рода не фиксируют в явном виде, выбирая лишь вероятность ошибки первого рода  $\alpha$  (или  $P$ ) и способ проверки гипотезы.

Вероятность ошибки первого рода при проверке гипотез часто называют также **уровнем значимости**.

Величину  $\bar{P} = 1 - \alpha$  называют **доверительной вероятностью**. Ее смысл в том, что опытные данные не противоречат основной гипотезе с вероятностью  $\bar{P} = 1 - \alpha$  ( $\bar{P} = 1 - P$ ). Часто величину доверительной вероятности выражают в процентах:  $\bar{P} = (1 - \alpha) \cdot 100\%$ .

Во многих случаях важную роль играет выбор альтернативной гипотезы (обычно обозначаемой как  $H_1$ ). Альтернативную основную гипотезе  $H_0$  гипотезу  $H_1$  выбирают таким образом, чтобы вместе они удовлетворяли следующим условиям:

- гипотезы  $H_0$  и  $H_1$  должны взаимно исключать друг друга;
- вместе гипотезы  $H_0$  и  $H_1$  должны покрывать все возможные варианты ответа.

Если возможны разные альтернативные гипотезы, выбор проводится исходя из свойств исследуемого объекта.

Классическим **примером** является сравнение математических ожиданий двух совокупностей результатов измерений. Основная гипотеза  $H_0$ : математические ожидания совокупностей, из которых получены выборки, равны. Альтернативная гипотеза  $H_1$  может быть сформулирована двояко:

- математические ожидания совокупностей, из которых получены выборки, не равны; либо:
- математическое ожидание одной из совокупностей больше математического ожидания другой.

Если из априорных соображений известно, что одно из математических ожиданий заведомо больше другого, то выбирается вторая гипотеза, а в отсутствие такой информации — первая. Способ проверки основной гипотезы (одно- или двухсторонний статистический тест) зависит от того, какая альтернативная гипотеза выбрана. Подробнее см. разд. 2.5.3.

Отметим, что в результате проверки **гипотезы либо отвергаются, либо не отвергаются**; неправильно говорить, что какая-либо гипотеза принимается или верна. Эта терминология отражает тот факт, что проверка гипотез всегда проводится на основе неполной информации.

Ситуация, когда информации достаточно для отклонения гипотезы с принятой доверительной вероятностью, обычна. Решение о том, что гипотеза должна быть отвергнута с заданной вероятностью вряд ли изменится при проведении дополнительных исследований (при увеличении размеров выборки).

Наоборот, принимая гипотезу, т. е. на основании полученных экспериментальных данных утверждая, что она верна, мы можем допустить грубую ошибку — на самом деле гипотеза неверна, просто в нашем распоряжении недостаточно данных,



чтобы ее отвергнуть с принятым уровнем доверительной вероятности. Очень часто гипотеза отвергается после проведения дополнительных измерений.

Выделяют два основных способа проверки гипотез:

- с использованием подходящего статистического теста, при этом рассчитанное по результатам измерений значение статистической функции (выбор функции и способ расчета зависят от проверяемой гипотезы) сравнивают с табличным значением при заранее выбранной доверительной вероятности (многочисленные примеры использования этого способа описаны ниже);
- расчет доверительного интервала (интервалов), соответствующего выбранной доверительной вероятности, и проверка того, попадает ли искомое значение в этот интервал (пересекаются ли интервалы). Такой способ часто применяется при проверке гипотез о равенстве двух величин. Пример использования доверительных интервалов при проверке гипотез см. в разд. 2.5.3.

С теоретической точки зрения оба способа эквивалентны и приводят к одинаковым выводам. Выбор одного из них определяется в основном из соображений удобства расчетов.

Важнейшим вопросом при проверке гипотез является выбор доверительной вероятности ошибки первого рода или, что фактически то же самое, доверительной вероятности  $\bar{P}$ . Очень часто при обработке результатов измерений выбирают  $\bar{P} = 0,95$  для двухстороннего теста; можно сказать даже, что такая доверительная вероятность является стандартной. Доверительная вероятность означает, что ошибка первого рода может быть допущена не более чем в одном случае из двадцати ( $\alpha = 0,05$ ). Однако это не догма, и доверительная вероятность может выбираться другой в зависимости от последствий ошибочного решения. Так, при решении многих производственных задач часто выбирают  $\bar{P} = 0,90$ , так как ошибка может повлечь лишь небольшие финансовые потери. В случае же задач, связанных с оценкой токсического воздействия лекарств на организм человека, выбирают равной 0,99 или даже 0,999 — использование ядовитых лекарств недопустимо.

Многие компьютерные программы при проверке гипотез выводят непосредственно значение вероятности  $\alpha$  (обозначаемой также как  $P$ ). В этом случае интерпретация результатов проверки гипотез выглядит так: гипотеза отвергается, когда рассчитанная величина  $\alpha$  (она же  $P$ ) меньше заранее выбранного значения.

**Пример.** При проверке гипотезы о равенстве двух дисперсий исходно был принят уровень доверительной вероятности  $\bar{P} = 0,99$  (то есть  $P = 0,01$ ). При расчетах с помощью компьютерной программы SPSS вероятность ошибки первого рода  $P$  оказалась равной  $P = 0,006281$ . Гипотеза отвергается с принятым уровнем доверительной вероятности, так как  $0,006281 < 0,01$ .

## 2.2. Способы изображения и характеристики выборок

В простейшем случае мы имеем результаты нескольких измерений одной и той же величины  $X_1, X_2, \dots, X_p, \dots, X_N$ , которую согласно основному постулату математической статистики можно рассматривать как *случайную выборку* из некоторой совокупности. Результаты измерений  $X_i$  называют также *вариантами*. Часто при выполнении расчетов результаты измерений (варианты) располагают по возрастанию

или по убыванию величин  $X_i$ . В этом случае говорят об *упорядоченной выборке*. При расположении по возрастанию величины выборка записывается так:

$$X_1 \leq X_2 \leq \dots \leq X_i \leq \dots \leq X_N \quad (2.5)$$

Ниже, когда говорится об упорядоченной выборке, всегда имеется в виду выборка, упорядоченная по возрастанию, т. е. вида (2.5), если иное не оговорено особо.

### 2.2.1. Изображение выборок

Для наглядного представления выборок (эмпирических распределений) используются два основных способа.

*Эмпирическая функция распределения, или полигон частот*, строится следующим образом. Для каждого действительного числа  $X$  подсчитывают число выборочных значений  $m_i(X)$ , меньших  $X$ . Ступенчатая функция  $F_n(X) = m_i(X)/N$  и есть эмпирическая функция распределения выборки (рис. 2.3). При  $N \rightarrow \infty$  эмпирическая функция распределения почти наверное сходится к функции распределения  $\Phi(X)$  генеральной совокупности (см. разд. 2.1.1).

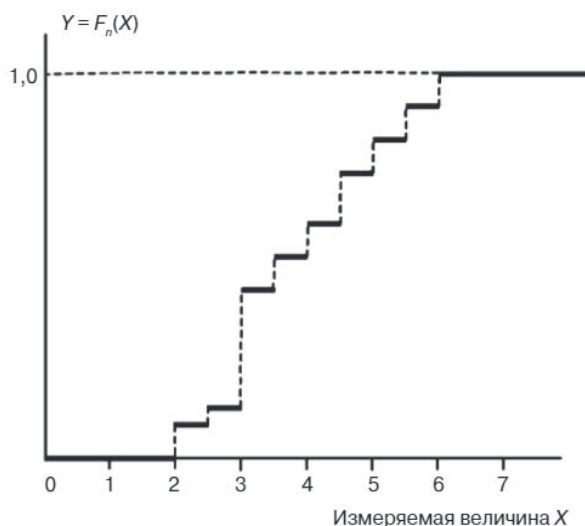


Рис. 2.3. Выборочная функция распределения (полигон частот),  $N = 9$

Более наглядным и чаще используемым способом графического изображения, представляющим распределение величин  $X_j$ , является гистограмма, или столбчатая диаграмма. Для ее построения диапазон, в котором расположены величины  $X_j$ , разбивают на одинаковые интервалы или классы (равномерное разбиение). Обозначим число результатов измерений  $X_j$ , попавших в каждый  $j$ -й интервал, как  $n_j$  («групповые частоты»), а число интервалов — как  $j$ . При построении гистограммы над каждым  $j$ -м интервалом чертится прямоугольник, высота которого равна соответствующей групповой частоте, деленной на общее число измерений  $N$ , то есть  $n_j/N$ . Пример гистограммы показан на рис. 2.4.

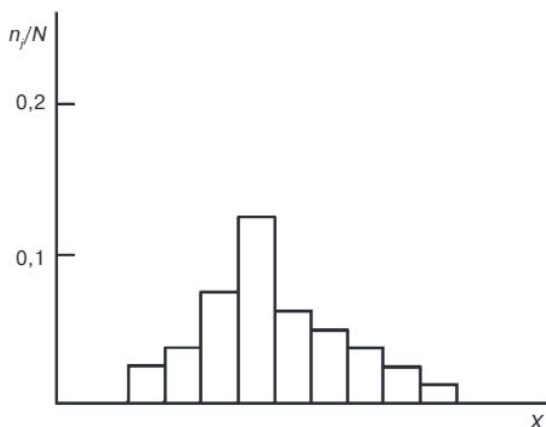


Рис. 2.4. Гистограмма (столбчатая диаграмма)

Гистограмму важно построить так, чтобы, рассматривая ее, получить максимум информации. Основная проблема — выбор числа интервалов  $j$  (или, что то же самое, их размера), от чего зависит вид гистограммы. Не существует правила оптимального разбиения: построение удачной гистограммы все еще остается предметом интуиции и искусства исследователя. При малом числе интервалов разбиения (интервал велик) плохо выявляются характерные особенности распределения. С ростом числа интервалов характерные особенности выявляются лучше. Однако при слишком большом числе интервалов (интервал слишком мал) гистограмма снова теряет характерные особенности распределения, превращаясь в пределе (когда в каждом интервале не более одного значения  $X_i$ ) в чередование «пустых» интервалов и одинаковых по высоте прямоугольников.

Оптимальное число интервалов при  $N < 70$  можно оценить по формуле  $j \approx \sqrt{N}$ . При больших  $N$  оптимальное число интервалов растет медленнее, чем  $\sqrt{N}$ , и меняется в пределах 9–11 для  $N < 1000$ . Более точно оптимальное число интервалов можно оценить по формуле  $j \approx 1 + 3,332 \lg N$ .

Отметим, что при  $N \rightarrow \infty$  огибающая гистограммы почти наверное переходит в функцию плотности распределения генеральной совокупности  $\varphi(X)$  (см. разд. 2.1.1).

По виду гистограммы легко судить о многих характеристиках эмпирического распределения (количество максимумов, асимметрия и т. д.).

Эмпирическая функция распределения и гистограмма наглядны при достаточно большом числе измерений ( $N > 30$ ).

### 2.2.2. Статистические характеристики выборок

Менее наглядным, но зато количественным способом охарактеризовать выборку является расчет *статистических показателей (характеристик)* этой выборки, которые являются оценками соответствующих параметров генеральной совокупности. Чаще всего используют следующие статистические показатели.

1. *Среднее арифметическое значение* (часто называемое средним значением или просто средним) выборки вида (2.1 или 2.5)



$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i. \quad (2.6)$$

При достаточно большом  $N$  и в отсутствие резко выделяющихся значений результатов измерений (грубых промахов, см. разд. 2.4.2) среднее арифметическое является хорошей оценкой математического ожидания генеральной совокупности<sup>18</sup>. Можно показать, что среднее значение выборки  $\bar{X}$  является несмещенной оценкой математического ожидания  $\mu$ , соответствующей совокупности<sup>19</sup>.

2. **Медиана, или срединное значение (median)** для упорядоченной выборки (2.5) есть:

- при нечетных  $N$  – такое значение  $X_p$ , число результатов измерений справа и слева от которого одинаково;
- при четных  $N$  – среднее арифметическое двух срединных членов последовательности, справа и слева от которых лежит одинаковое число результатов измерений.

**Примеры.** Медиана результатов измерений 1, 1; 1, 2; 1, 5 есть 1, 2. Медиана выборки 12, 4; 12, 4; 12, 6; 15, 2 равна 12, 5.

Медиана, как и среднее значение, является оценкой математического ожидания генеральной совокупности. Она малочувствительна к наличию грубых промахов. Так, для последней выборки медиана не зависит от явно выделяющегося результата 15, 2. Медиану часто используют вместо среднего арифметического для характеристики малых ( $N < 8$ ) выборок.

3. **Стандартное отклонение  $s$**  выборки, рассчитываемое по формуле

$$s = +\sqrt{\sum_{i=1}^N \frac{(X_i - \bar{X})^2}{N-1}} = +\sqrt{\sum_{i=1}^N X_i^2 - N\bar{X}^2}, \quad (2.7)$$

где  $N - 1$  есть число степеней свободы.

Величина  $s^2$  называется **дисперсией**.

Стандартное отклонение, как и дисперсия, является мерой рассеяния (разброса) результатов измерений.

Стандартное отклонение  $s$  представляет собой смещенную оценку соответствующего параметра  $\sigma$  совокупности (в отличие от  $s^2$ , являющейся несмещенной оценкой  $\sigma^2$ ). Несмещенная оценка может быть получена делением величины  $s$ , рассчитанной по формуле (2.7), на так называемый **момент отношения  $s/\sigma$** , обозначаемый как  $M_v$  (см. табл. 2.1).

<sup>18</sup> Иногда, если распределение результатов близко к логарифмически нормальному, вместо среднего арифметического используют среднее геометрическое значение.

<sup>19</sup> Для обозначения параметров выборки (оценок) мы используем латинские буквы ( $\bar{X}$ ,  $s$ ,  $A$ ,  $E$ ,  $r$  и т. д.), а для соответствующих параметров генеральных совокупностей – греческие ( $\mu$ ,  $\sigma$ ,  $\gamma_1$ ,  $\gamma_2$ ,  $\rho$  и т. д.). Такой подход является общепринятым.

Таблица 2.1. Величины  $M_v$  (моменты отношения  $s/\sigma$ ) в зависимости от объема выборки  $N$  [23]<sup>20</sup>

$N$	2	3	4	5	6	7	8	9	10
$M_v$	0,798	0,886	0,921	0,940	0,951	0,959	0,965	0,969	0,973

Видно, что при увеличении  $N$  величина  $M_v$  быстро стремится к 1. На практике обычно считают  $s$  несмещенной оценкой для  $\sigma$  при  $N \geq 8$ .

Часто используют также *относительное стандартное отклонение* (о.с.о. или *RSD*, от англ. *Relative Standard Deviation*)<sup>21</sup>, представляющее собой отношение стандартного отклонения выборки к ее среднему значению:

$$RSD = \frac{s}{\bar{X}}. \quad (2.8)$$

Стандартное отклонение и среднее значение выборки  $\bar{X}$  независимы тогда и только тогда, когда выборка получена из нормальной совокупности.

При определении стандартного отклонения мы ввели понятие «число степеней свободы» — одно из основных в математической статистике. Это понятие трудно поддается строгому определению. В простых случаях число степеней свободы может быть интерпретировано как число переменных, которые могут быть произвольно присвоены при расчете той или иной статистики<sup>22</sup>. Можно сказать также, что это — число независимых случайных величин, по которым вычисляется та или иная статистика или количество значений в итоговом вычислении статистики, способных варьироваться. Число степеней свободы есть разность между числом измерений и количеством коэффициентов модели, вычисленных по результатам этих экспериментов независимо друг от друга.

Другими словами, для выборки простейшей структуры (2.1) число степеней свободы равно разности между числом измерений  $N$  и числом параметров, зафиксированных при расчете данной статистики. Для стандартного отклонения  $s$  число измерений равно  $N$ , при ее расчете зафиксирован один параметр (среднее значение  $\bar{X}$ ) и, следовательно, число степеней свободы равно  $N - 1$ .

4. Размах  $R$  для упорядоченного по возрастанию  $X_i$  ряда (2.5), равен

$$R = X_N - X_1, \quad (2.9)$$

<sup>20</sup> В оригинальной работе индекс  $v$  есть число степеней свободы. Поскольку при оценке  $s$  мы обычно не знаем математического ожидания величины  $X$ ,  $v = N - 1$ .

<sup>21</sup> Относительное стандартное отклонение часто обозначают также как  $s_r$ . Мы не будем использовать это обозначение, чтобы не путать относительное стандартное отклонение, рассчитываемое по формуле (2.8), и стандартное отклонение повторяемости (сходимости) — см. разд. 3.3.1 и сл.

<sup>22</sup> Термин «статистика» здесь использован в узком смысле — это любая подходящая функция  $X = f(X_1, X_2, \dots, X_r, \dots, X_N)$ , значение которой может быть предсказано с большей точностью, чем единичные результаты измерений  $X_i$ . Такое свойство функции  $X$  называется «статистической устойчивостью». Из этого определения вытекает, что рассматриваемые нами оценки  $\bar{X}$ ,  $s$ ,  $A$ ,  $E$ ,  $r$  и другие являются статистиками.

т. е. размах есть разность максимального и минимального результатов измерений в выборке. Размах часто используется вместо стандартного отклонения для характеристики разброса результатов измерений при малых выборках ( $N < 8$ ). Для таких выборок использование стандартного отклонения менее эффективно.

5. *Асимметрия (выборочный коэффициент асимметрии)  $A$* , рассчитываемая по формуле

$$A = \frac{\sum_{i=1}^N (X_i - \bar{X})^3}{N \cdot s^3}. \quad (2.10)$$

Эта характеристика, как следует из ее названия, отражает асимметрию распределения. Для симметричных распределений она равна нулю. Для распределений, вытянутых (имеющих «хвост») в сторону больших значений,  $A > 0$ , для имеющих «хвост» в сторону малых значений  $A < 0$ . Положительная асимметрия ( $A > 0$ ) в задачах, связанных с измерениями, встречается чаще. Так, положительную асимметрию обычно имеют распределения результатов измерения концентрации веществ, близкой к пределу обнаружения метода анализа.

6. *Экссесс (или выборочный коэффициент эксцесса)  $E$* , рассчитываемый по формуле

$$E = \frac{\sum_{i=1}^N (X_i - \bar{X})^4}{N \cdot s^4}. \quad (2.11)$$

Экссесс, рассчитываемый по формуле (2.11), близок к нулю для выборок из генеральной совокупности с нормальным распределением. Положительный эксцесс соответствует распределению более островершинному, чем нормальное, отрицательный — наоборот. Значительный (отрицательный или положительный) эксцесс обычно возникает при нарушении предположения о чисто случайном характере распределения, в ситуациях, когда рассматриваемое распределение на самом деле является суммой двух или нескольких распределений.

*Кроме перечисленных характеристик, используются и другие.*

*Квантиль  $Q(p)$   $p$ -го порядка* делит генеральную совокупность на две части в пропорции  $p$ :  $(1 - p)$ . Если квантиль применяется для характеристики упорядоченной выборки вида (2.5), то:

- если  $N/(p + 1)$  есть целое число, то квантиль  $Q(p)$  равен  $X_i = N/(p + 1)$ ;
- если  $N/(p + 1)$  — дробное число, расположенное между  $X_i$  и  $X_{i+1}$ , то  $Q(p)$  — величина, равная  $(X_i + X_{i+1})/2$ . В последнем случае в качестве квантиля используют также величину  $p(X_i + X_{i+1})/(1 - p)$ . Квантиль порядка  $p = 1/2$  есть медиана выборки.

Для упорядоченной по возрастанию выборки вида (2.5) порядковый номер результата измерений (варианты)  $X_i$  называют также ее *рангом  $p(i)$* . Величина  $p(i) = (i - 0,5)/N$  называется *скорректированным ранговым отношением*, или *ранговым отношением*. Очевидно,  $i$ -я варианта упорядоченной выборки делит ее в отношении  $p(i)$ :  $(1 - p(i))$  и  $i$ -я варианта есть  $Q(p(i))$  квантиль выборки.



## 2.3. Теоретические распределения

Как правило, выборку оценивают с точки зрения близости распределения соответствующей ей совокупности тому или иному *теоретическому распределению* (теоретической функции распределения).

В математической статистике рассматриваются как теоретические распределения, призванные описывать реальные результаты измерений, так и распределения, используемые в «служебных» целях, например, в ходе проверки гипотез (специальные распределения). Рассмотрим некоторые из них.

### 2.3.1. Непрерывные распределения

Случайная величина называется непрерывной, если ее функция распределения  $\Phi(X) \equiv P(\xi < X)$  непрерывна и имеет кусочно-непрерывную производную – плотность непрерывного распределения  $\varphi(X)$ . Далее вместо «функция распределения случайной величины» говорится просто «функция распределения».

Непрерывные распределения обычно используются при описании результатов измерений с большим числом возможных исходов. Например, при измерении массы на аналитических весах результат может лежать в интервале от 0 до 100 г с шагом 0,0001 г, т. е. возможны  $\sim 10^6$  исходов; при измерении оптической плотности на спектрофотометре результат может лежать в интервале 0–2 с шагом 0,001, т. е. возможны  $\sim 2000$  исходов. В то же время бывают измерения, при которых может реализоваться всего несколько (до 10–20) исходов. К таким результатам применяют дискретные распределения (см. ниже, разд. 2.3.2).

Важнейшим непрерывным распределением является *нормальное распределение*.

#### 2.3.1.1. Нормальное распределение

При описании результатов измерений методами математической статистики чаще всего используется *нормальное распределение (распределение Гаусса)*<sup>23</sup>.

Причины того, что нормальному распределению придается столь важное значение (а оно даже изображалось на некоторых купюрах евро) следующие. Согласно центральной предельной теореме, если случайную величину можно представить как сумму большого числа не зависящих друг от друга слагаемых, каждая из которых вносит в сумму незначительный вклад, то распределение этой суммы близко к нормальному. Случайные погрешности измерений (см. разд. 1.3.5) как раз и определяются большим числом факторов, каждый из которых незначителен по сравнению с другими. Эта ситуация рассматривается центральной предельной теоремой, и, следовательно, *случайные погрешности измерений должны иметь приблизительно нормальное распределение*. Именно так и обстоит дело на практике.

Кроме того, нормальное и связанные с ним распределения обладают целым рядом уникальных математических свойств, делающих возможным решение с их использованием многих статистических задач.

<sup>23</sup> Мы рассматриваем главным образом одномерные распределения (эмпирические и теоретические).

Случайная величина называется распределенной нормально, если плотность ее распределения описывается формулой

$$h(X) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(X-\mu)^2}{2\sigma^2}\right], \quad (2.12)$$

где  $\mu$  (математическое ожидание) и  $\sigma$  (стандартное отклонение) – параметры распределения.

Если случайная величина  $X$  имеет нормальное распределение с параметрами  $\mu$  и  $\sigma$ , то часто говорят, что она «распределена нормально». Этот факт кратко записывается как  $N(X; \mu, \sigma^2)$ . Другой вариант записи:  $X \in N(\mu, \sigma^2)$ . Иногда пишут «величина ... распределена нормально  $N(\mu, \sigma^2)$ ».

График функции (2.12) представляет собой колоколообразную кривую, причем параметр  $\mu$  есть точка максимума, через которую проходит ось симметрии, а параметр  $\sigma$  – расстояние от этой оси до каждой из точек перегиба. При уменьшении  $\sigma$  распределение становится более узким и острым, а при увеличении – уширяется и становится более плоским.

Рассчитывать значение функции (2.10) для каждого конкретного сочетания  $\mu$  и  $\sigma$  «с нуля» неудобно. На практике исходят из нормированного и центрированного нормального распределения  $X \in N(0, 1)$ , то есть нормального распределения с параметрами  $\mu = 0$  и  $\sigma^2 = 1$ .

Плотность такого распределения

$$\varphi(X) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} X^2\right] \quad (2.13)$$

и соответствующее интегральное распределение (гауссов интеграл ошибок)

$$\Phi(X) = \int_{-\infty}^X \varphi(t) dt \quad (2.14)$$

табулированы [23], а также могут быть рассчитаны для любого значения  $X$  с помощью компьютерных программ.

График плотности распределения (2.13) показан на рис. 2.2.

Зная значение плотности нормального распределения  $X \in N(0, 1)$  плотность нормального распределения с параметрами  $(\mu, \sigma^2)$  легко найти из соотношения

$$N(X; \mu, \sigma) = \frac{1}{\sigma} N\left(\frac{X-\mu}{\sigma}; 0, 1\right). \quad (2.15)$$

### 2.3.1.2. Моменты непрерывного распределения

Любое непрерывное распределение характеризуется рядом параметров, называемых **моментами распределения**.

Предположим, мы имеем непрерывную случайную величину  $X$  с функцией распределения  $\Phi(X)$  и плотностью  $\varphi(X)$ . Тогда  **$i$ -м начальным моментом** называется величина  $v_i$

$$v_i = \int_{-\infty}^{+\infty} X^i \varphi(X) dt \quad (i=1, 2, \dots), \quad (2.16)$$



а  $i$ -м центральным моментом — величина  $\mu_i$

$$\mu_i = \int_{-\infty}^{+\infty} X^i (X - v_i)^i \Phi(X) dX \quad (i = 1, 2, \dots). \quad (2.17)$$

Первый начальный момент ( $i = 1$ ) называется *математическим ожиданием* случайной величины  $X$  и обозначается как  $M(X)$ . Для нормального распределения это — параметр распределения  $\mu$ . Оценками математического ожидания генеральной совокупности являются среднее арифметическое значение и медиана выборок из этой совокупности.

Если функция распределения описывает нормальную совокупность результатов измерений, то *(в отсутствие систематических погрешностей) математическое ожидание есть истинное значение измеряемой величины*.

Второй центральный момент  $\mu_2$  называется дисперсией случайной величины  $X$  и обозначается как  $DX$ , или  $\sigma^2$ . Величина  $\sigma = +\sqrt{DX}$  называется *стандартным отклонением*, или *средним квадратичным отклонением* величины  $X$ . Для нормального распределения эта величина является параметром распределения  $\sigma$ . Оценками стандартного отклонения и дисперсии генеральной совокупности являются стандартное отклонение и дисперсия выборок из этой совокупности.

Третий и четвертый центральные моменты совокупности называются, соответственно, *коэффициент асимметрии*

$$\gamma_1 = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}} \quad (2.18)$$

и *коэффициент эксцесса*

$$\gamma_2 = \frac{\mu_4}{(\mu_2)^2} - 3. \quad (2.19)$$

Цифра 3 в (2.19) и в (2.11) вычитается для удобства, поскольку отношение  $\mu_4/(\mu_2)^2$  для нормального распределения равно 3. Таким образом, величины  $\gamma_1$  и  $\gamma_2$  для нормального распределения равны нулю.

Оценками коэффициентов асимметрии и эксцесса генеральной совокупности являются асимметрия и эксцесс, рассчитанные для выборок из этой совокупности (см. выше, разд. 2.1.2).

Моменты более высоких порядков и связанные с ними статистические показатели трудно поддаются наглядной интерпретации и практически не используются.

Приведем еще две характеристики, применяемые при совместном описании двух случайных величин. Ковариация (смешанный момент второго порядка) двух случайных величины  $X$  и  $Y$ , имеющих математические ожидания  $M(X)$  и  $M(Y)$ , есть величина

$$\sigma_{XY} = M[X - M(X)][Y - M(Y)]. \quad (2.20)$$

*Коэффициент корреляции*  $\rho_{XY}$  между теми же случайными величинами определяется как

$$\rho_{XY} = M \left( \frac{X - M(X)}{\sigma_X} \times \frac{Y - M(Y)}{\sigma_Y} \right), \quad (2.21)$$

где  $\sigma_X$  и  $\sigma_Y$  — стандартные отклонения случайных величин  $X$  и  $Y$ . Эти параметры характеризуют взаимосвязь величин  $X$  и  $Y$  (подробнее см. разд. 2.7.1).

### 2.3.1.3. Равномерное и треугольное распределения

**Равномерное распределение.** Случайная величина *равномерно распределена* на отрезке  $[a, b]$ , если ее плотность распределения постоянна на этом отрезке и равна нулю во всех остальных точках числовой оси. График плотности равномерно распределенной случайной величины показан на рис. 2.5. Интегральная функция такого распределения равна:

$$\begin{cases} 0 & \text{при } X < a; \\ (X - a) / (b - a) & \text{в интервале } a < X < b; \\ 1 & \text{при } X > b. \end{cases}$$

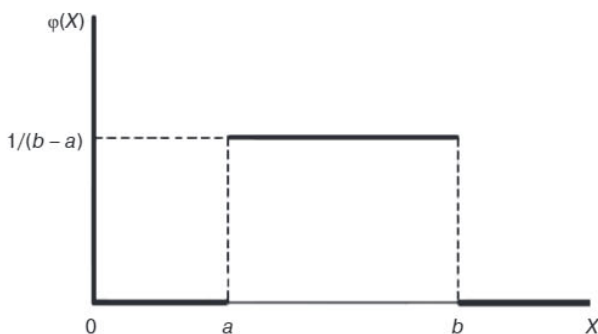


Рис. 2.5. График плотности распределения случайной величины, равномерно распределенной на отрезке  $(a, b)$

Дисперсия равномерно распределенной случайной величины равна

$$\sigma^2 = (b - a)^2 / 12. \quad (2.22)$$

Если обозначить отрезок  $[a, b]$  как  $(\bar{X} \pm \Delta)$ , где  $\bar{X} = (a + b) / 2$  и полуширина отрезка  $\Delta = |a - b| / 2$ , то (2.22) можно записать как

$$\sigma^2 = \frac{\Delta^2}{3}. \quad (2.23)$$

Равномерное распределение на некотором отрезке часто приписывают случайным величинам, относительно которых известно только, что их математическое ожидание лежит где-то внутри этого отрезка. Так обстоит дело, в частности, со стандартными образцами — известно, что истинное значение концентрации в образце лежит в некотором интервале, и не более того. Приходится полагать распределение концентрации в этих образцах равномерным (подробнее см. разд. 3.5.2.3.2).

**Треугольное распределение (распределение Симпсона).** Случайная величина имеет треугольное распределение на отрезке  $[a, b]$  ( $a < b$ ), если ее функция распределения равна

$$\begin{cases} \frac{2}{b-a} - \frac{2}{(b-a)^2} |a+b-2X|, & \text{если } X \in [a, b] \\ 0, & \text{если } X \notin [a, b] \end{cases}$$

Функция плотности распределения такой случайной величины показана на рис. 2.6.

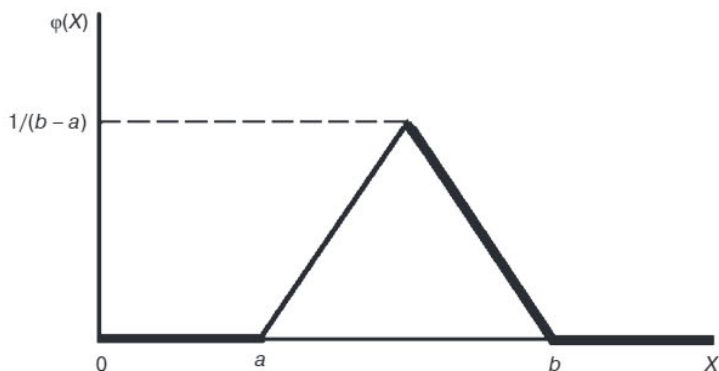


Рис. 2.6. График плотности распределения случайной величины, имеющей треугольное распределение (распределения Симпсона) на отрезке  $(a, b)$

Дисперсия случайной величины, имеющей треугольное распределение, равна

$$\sigma^2 = (b - a)^2 / 24. \quad (2.24)$$

Обозначив отрезок  $[a, b]$  как  $(\bar{X} \pm \Delta)$ , где  $\bar{X} = (a + b) / 2$  и  $\Delta = |a - b| / 2$  — полуширина отрезка, получаем

$$\sigma^2 = \frac{\Delta^2}{6}. \quad (2.25)$$

Треугольное распределение на некотором отрезке часто приписывают случайным величинам, функция распределения которых неизвестна, но есть информация, что математическое ожидание лежит ближе к середине отрезка, чем к краям. Такое распределение используют, например, для описания распределения объемов, отмеряемых мерными колбами или пипетками, для которых указывается, что дозируемый объем лежит в некотором интервале. Это связано с тем, что при их калибровке истинное значение лежит ближе к середине этого интервала (см. разд. 2.5.3.3.2).

В практике встречаются также измерения, погрешности которых лучше описываются другими непрерывными распределениями: логарифмически нормальным, распределением Коши и т. д. Сводка различных распределений дана, например, в [3].

#### 2.3.1.4. Некоторые специальные распределения

Существует несколько непрерывных распределений, связанных с нормальным распределением и играющих важную роль при решении задач математической статистики (нахождение доверительных интервалов, проверка гипотез). Чаще всего используются  $\chi^2$ -распределение (распределение Пирсона), распределение Стюдента, называемое также  $t$ -распределением и  $F$ -распределение (распределение Фишера).

**$\chi^2$ -распределение (Пирсона)** с  $n$  степенями свободы есть распределение  $\chi_n^2$  случайной величины, представляющей собой сумму квадратов  $n$  взаимно независимых случайных величин  $\xi_1, \xi_2, \dots, \xi_n$ , каждая из которых нормально распределена с параметрами  $(0, 1)$ .

Очевидно, что  $\chi_n^2$  – распределение для простоты можно рассматривать как совокупность распределений, каждое из которых соответствует какому-либо числу степеней свободы  $n$ . Функция распределения для  $\chi^2$  расположена в первом квадранте и стремится к нормальному распределению при  $n \rightarrow \infty$ . Значения  $\chi^2$  находят из таблиц (см. Приложение 3) или рассчитывают с помощью компьютерных программ. При проверке гипотез часто используется и другой подход: компьютерная программа рассчитывает вероятность  $P$  по значению  $\chi^2$  и числу степеней свободы  $n$ . Одно из применений  $\chi^2$ -критерия – проверка гипотезы о форме распределения случайной величины по критерию Пирсона – описано в разд. 2.4.1.

**Распределение Стюдента ( $t$ -распределение,  $t$ -распределение Стюдента)** описывает распределение средних значений результатов измерений при небольших выборках. Случайной величиной, подчиняющейся распределению Стюдента с  $f = N - 1$  степенями свободы, называется следующее отношение независимых случайных величин

$$t_N = \frac{\xi}{\sqrt{\frac{1}{f} \chi^2 f}}, \quad (2.26)$$

где  $\xi$  – случайная величина, распределенная нормально с параметрами  $(0, 1)$ ,  $\xi \in N(0, 1)$ , а  $\chi^2$  подчиняется распределению  $\chi^2$  с  $f$  степенями свободы.

Для простоты  $t$ -распределение можно рассматривать как совокупность распределений, каждое из которых соответствует какому-либо числу степеней свободы  $f$ . Как указано выше, распределение Стюдента используют при работе с небольшими ( $N < 30$ ) выборками из нормально распределенной генеральной совокупности.

Абсцисса максимума плотности  $t$ -распределения совпадает с абсциссой максимума плотности нормального распределения, однако *ширина плотности распределения Стюдента больше ширины плотности нормированного нормального распределения, и различие тем больше, чем меньше число степеней свободы  $f$* . В случае больших выборок ( $N > 30$ ) вместо распределения Стюдента используют нормальное распределение (*при  $N \rightarrow \infty$  распределение Стюдента есть нормальное распределение с параметрами  $(0, 1)$* ).

В практике обработки результатов измерений используют таблицы (см. Приложение 4), в которых даны критические значения (процентные точки) нормированного распределения Стюдента, представляющие собой пределы интегрирова-



ния этого распределения в зависимости от *доверительной вероятности*  $\bar{P}$  (см. разд. 2.2.2) и числа степеней свободы  $f$ . Эти же величины можно рассчитать с помощью статистических компьютерных программ. Если мы проведем бесконечное число опытов, в каждом из которых будем получать выборку объемом  $N$  измерений, и для каждой из них рассчитывать среднее значение, то доля средних значений, попавших внутрь этих пределов, составит  $\bar{P}$ . Распределение Стьюдента используется при проверке гипотез, связанных с различием математических ожиданий, и при нахождении границ доверительных интервалов для средних значений (см. разд. 2.5.3 и сл.).

***F-распределение (распределение Фишера).*** Возьмем две независимые выборки  $X_i$  ( $i = 1, \dots, N_1$ ) и  $Y_j$  ( $j = 1, \dots, N_2$ ) из одной и той же совокупности, нормально распределенной с параметрами  $(0, 1)$ .  $F$ -распределением со степенями свободы  $f_1 = N_1 - 1$  и  $f_2 = N_2 - 1$  называется распределение величины

$$F = \frac{\sum_{i=1}^{N_1} (X_i - \bar{X})^2 / (N_1 - 1)}{\sum_{j=1}^{N_2} (Y_j - \bar{Y})^2 / (N_2 - 1)} = \frac{s_X^2}{s_Y^2}, \quad (2.27)$$

где  $s_X^2$  и  $s_Y^2$  — дисперсии выборок. Отношение (2.27) обычно строится таким образом, чтобы выполнялось соотношение  $F \geq 1$ , т. е. числитель не должен быть меньше знаменателя<sup>24</sup>. Аналогично тому, как это имеет место для  $\chi^2$ -распределения и  $t$ -распределения Стьюдента, можно сказать, что  $F$ -распределение представляет собой множество распределений, каждое из которых соответствует одному из сочетаний чисел степеней свободы  $(f_1, f_2)$ .

$F$ -распределение применяется при проверке гипотезы о равенстве двух дисперсий (разд. 2.5.1), в том числе и при дисперсионном анализе (разд. 2.6.2 и сл.). В практике измерений используют двумерную таблицу (Приложение 5), в которой даны процентные точки  $F$ -распределения, т. е. пределы интегрирования этого распределения в зависимости от доверительной вероятности  $\bar{P}$  и чисел степеней свободы  $f_1$  и  $f_2$ ; эти же величины можно рассчитать с помощью статистических компьютерных программ.

### 2.3.2. Распределения дискретных случайных величин.

#### *Распределение Пуассона*

Все рассмотренные выше распределения были непрерывными.

Если случайная величина не является непрерывной, а может принимать только некоторые фиксированные значения, то говорят о дискретной случайной величине и *дискретном распределении*.

Вообще говоря, при измерениях во всех случаях имеют дело с дискретными случайными величинами, поскольку число возможных результатов всегда ограничено. Это обусловлено тем, что в ходе измерений считывание результата осуществляется

<sup>24</sup> Иногда встречается обратная ситуация. Например, некоторые компьютерные программы рассчитывают величину  $F < 1$  — отношение меньшей дисперсии к большей.

с конечной точностью, определяемой ценой деления шкалы измерительного прибора (округление).

**Пример.** При измерении напряжения цифровым вольтметром в диапазоне 0–10 В и цене деления 0,01 В (максимально) возможное число исходов составляет 1000, а результаты 0,2578 В или 6,331 невозможны.

Если число возможных результатов измерения достаточно велико (как в рассматриваемом примере), этим обстоятельством пренебрегают и считают измеряемую величину непрерывной. В противном случае использование такого приближения становится некорректным (обычно при числе возможных исходов не более 15–20) и рассматривают дискретные распределения.

**Пример.** При измерении концентрации веществ радиометрическими методами анализа — при небольшой радиоактивности пробы число распадов, регистрируемых за время измерения, из соображений безопасности обычно невелико — в диапазоне от 6 до 15 и не может быть большим.

В таких случаях для описания результатов измерений в качестве теоретического распределения наиболее часто используется *распределение Пуассона*, которое выведено для случая, когда генеральная совокупность не изменяется в ходе эксперимента. В описанном выше примере, как и при большинстве измерений с дискретным распределением совокупностей, это предположение хорошо выполняется: общее количество радиоактивных атомов велико, и изменением его при распаде нескольких атомов можно пренебречь.

Распределение Пуассона имеет вид:

$$P = \frac{\mu^X e^{-\mu}}{X!}, \quad (2.28)$$

где  $P$  — вероятность появления значения  $X$ ,  $\mu$  — среднее значение. Интересно, что распределение Пуассона имеет только один параметр — среднее значение  $\mu$ ; стандартное отклонение распределения Пуассона со средним значением  $\mu$  равно  $\sigma = \sqrt{\mu}$ .

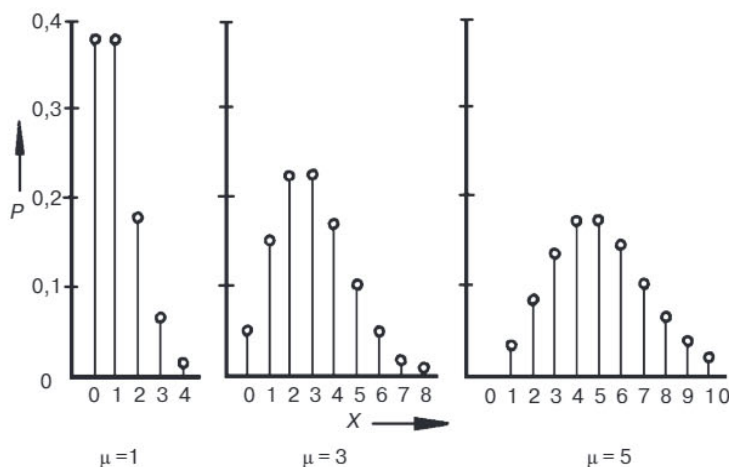


Рис. 2.7. Плотность распределения Пуассона для различных значений среднего арифметического  $\mu$

Распределение Пуассона (см. рис. 2.7) асимметрично тем более, чем меньше  $\mu$ . С ростом  $\mu$  асимметрия уменьшается; при  $\mu > 15-20$  распределение Пуассона практически совпадает с нормальным распределением с параметрами  $(\mu, \sqrt{\mu})$ .

Распределение Пуассона табулировано (см., например, [23]). Его используют аналогично распределению Стьюдента при расчете доверительных интервалов и проверке гипотез при малом числе возможных исходов измерений.

## 2.4. Общие условия применения методов математической статистики и проверка выполнения этих условий

Формальное, бездумное применение методов математической статистики при обработке результатов измерений часто приводит к неправильным выводам. Прежде чем использовать тот или иной статистический метод, надо быть уверенным, что выполняются его основные предположения.

В основе каждого конкретного метода, естественно, лежат свои предположения. Однако есть и достаточно общие, универсальные предположения, выполнение которых необходимо для корректного использования большинства методов статистического анализа.

1. Все рассмотренные ниже статистические методы основаны на взаимной независимости результатов измерений, полученных в одних условиях. Однако иногда оказывается, что выборка не является случайной и результаты отдельных измерений взаимозависимы (коррелированы).

**Пример.** При химическом анализе часть анализируемой пробы может оставаться в дозирующем устройстве и влиять на результат анализа следующей пробы. Это бывает при использовании одной пипетки для дозирования разных проб (если ее недостаточно тщательно ополаскивают), либо в автоматическом пробоотборнике («автосемплере») современных приборов.

Подобные эффекты можно обнаружить с использованием так называемых LAG-диаграмм, в которых результаты измерений, согласно их временной последовательности, наносят в зависимости друг от друга (чаще всего — зависимость результата  $i$ -го измерения от результата  $(i + 1)$ -го [22]. Конечно, оптимально строить измерения таким образом, чтобы результаты их были независимы — в нашем примере это означает либо тщательное ополаскивание дозирующего устройства между пробами, либо (при ручной работе) даже использование разных пипеток.

Строго говоря, рассмотренные ниже методы математической статистики непригодны для обработки взаимно коррелированных результатов. Отметим, однако, что коррелированность больше влияет на отдельные измерения, чем на результаты их статистической обработки.

2. Применение методов математической статистики при решении задач, связанных с проведением измерений, в том числе в области количественного химического анализа, с формальной точки зрения сводится к получению оценок и проверке гипотез. В большинстве случаев используют параметрические методы анализа данных, пригодные для выборок из совокупностей с известным распределением. Однако



они корректны только при выполнении определенных предположений, в рамках которых будет проводиться статистический анализ, т. е. модель, описывающая распределение совокупности и ее изменение во времени, должна быть корректной.

Полная проверка пригодности той или иной статистической модели в каждом конкретном случае чаще всего невозможна, поскольку для этого необходим большой статистический материал. Например, проверить гипотезу о нормальности распределения, имея 3–5 результатов измерений, нельзя. Поэтому для подтверждения адекватности модели проводят специальные эксперименты (выборочные исследования) или анализируют данные, полученные в течение длительного времени (например, при внутрилабораторном контроле; см. гл. 4). Иногда, впрочем, объем данных позволяет проверить корректность модели (например, в случае достаточно масштабных межлабораторных экспериментов, см. разд. 3.8).

При выборе модели необходимо проверять: соответствие распределения совокупности выбранному теоретическому распределению (чаще всего нормальному); случайность (независимость) единичных измерений; правильность выбранного вида зависимости между двумя или несколькими величинами (например, при построении градуировочной зависимости) и т. д. При использовании наиболее эффективных параметрических методов необходимо знать (хотя бы приблизительно) форму распределения генеральной совокупности.

### 2.4.1. Проверка гипотезы о распределении генеральной совокупности

Для проверки гипотезы о том, какое (теоретическое) распределение имеет совокупность, из которой получают выборки при использовании конкретного метода измерений, всего необходимо иметь достаточно большую выборку. Например, для проверки нормальности распределения случайной величины желательно провести несколько десятков измерений в одних условиях. Получаем выборку, содержащую  $N$  результатов измерений простейшей структуры (2.1).

Существуют следующие основные способы проверки гипотезы о распределении генеральной совокупности.

*Графический способ с использованием вероятностной бумаги* основан на использовании специальной вероятностной бумаги, похожей на обычную миллиметровую, но отличающейся от последней масштабом по оси  $Y$ , искаженным таким образом, что график гауссова интеграла ошибок (2.14), нанесенный на вероятностную бумагу, оказывается прямой линией. Этот способ пригоден только для проверки гипотезы о нормальном распределении генеральной совокупности.

Для проверки близости распределения выборки к нормальному распределению сначала разбивают диапазон, в котором расположены результаты измерений, на интервалы и рассчитывают число значений, меньших каждой из границ между этими интервалами, аналогично тому, как это делается при построении эмпирической функции распределения (см. разд. 2.2).

Обозначим каждую границу между интервалами как  $X_j$ , а число значений, меньших  $X_j$ , как  $Y_j$ . Если отложить на вероятностной бумаге по оси  $X$  величины  $X_j$ , а по оси  $Y$  — долю значений, меньших  $Y_j$ , т. е.  $Y_j/\Sigma Y_j$ , то для случайной выборки из нормальной совокупности нанесенные точки расположатся вокруг некоторой прямой (рис. 2.8).



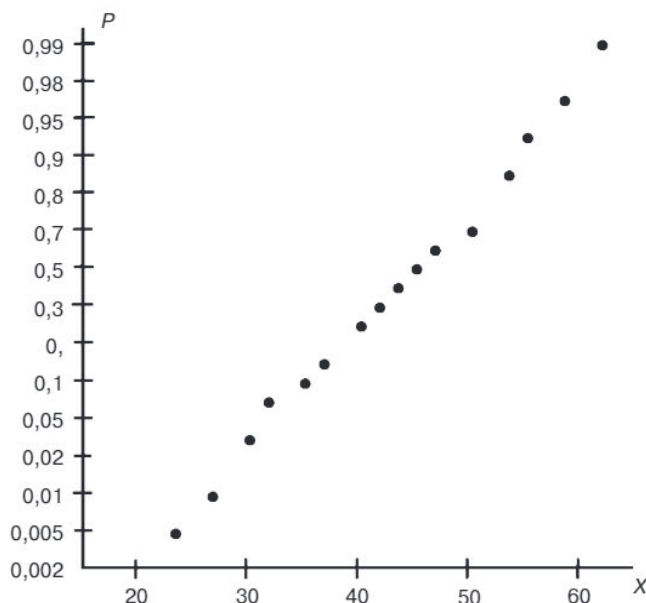


Рис. 2.8. Проверка нормальности распределения с помощью вероятностной бумаги

При этом на краях распределения (при накопленных частотах более 0,90 и менее 0,10) отклонения от прямой более вероятны, чем в середине, из-за меньшего числа точек в соответствующих интервалах.

Многие компьютерные программы выводят такое изображение на экран.

*Сравнение распределений с помощью Q–Q изображений* также относится к графическим методам. Этот способ имеет то преимущество перед способом с использованием вероятностной бумаги, что сравнивать можно распределения любой формы, как эмпирические, так и теоретические. Способ построения Q–Q изображения различен при сравнении двух эмпирических распределений и при сравнении эмпирического распределения с теоретическим.

Для сравнения двух эмпирических распределений равного объема (с одинаковым числом результатов измерений) изображают квантили одного из них в зависимости от квантилей другого распределения (рис. 2.9). Если эмпирические распределения являются выборками из одной и той же совокупности, то точки будут рассеяны вокруг биссектрисы первого координатного угла.

Для проверки гипотезы о близости эмпирического распределения теоретическому (т. е. распределению соответствующей совокупности) по одной оси откладывают квантили проверяемого эмпирического распределения, а по другой — квантили теоретического распределения. Как указано выше (см. разд. 2.1.2), квантиль теоретического распределения порядка  $P$  есть такое значение случайной величины  $X$ , для которого  $P\{X < p\} = P$ . Находят точки на координатной плоскости, соответствующие одинаковым квантилям. В случае близости эмпирического распределения теоретическому (что свидетельствует в пользу гипотезы о том, что совокупность имеет рассмотренное теоретическое распределение) точки располагаются вокруг прямой.

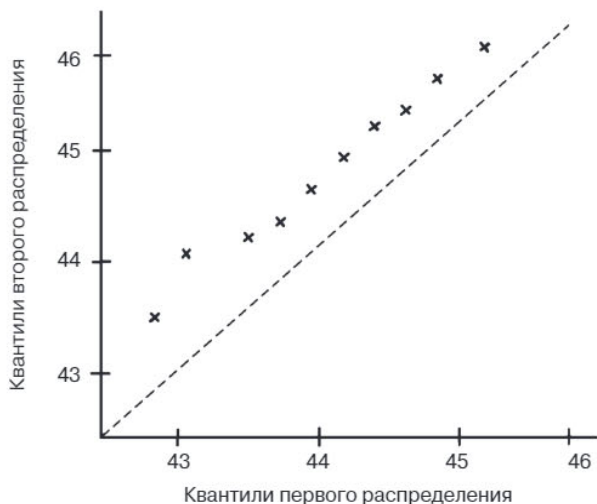


Рис. 2.9. Сравнение двух эмпирических распределений с равным числом результатов измерений (вариант) с помощью  $Q$ – $Q$  изображения. Показаны две выборки с близким стандартным отклонением, но разными средними значениями

Несколько примеров  $Q$ – $Q$  изображений выборок как из нормальных совокупностей, так и из совокупностей с другими распределениями даны в [22].

$Q$ – $Q$  изображения можно строить как с использованием статистических таблиц, так и компьютерным способом. Многие компьютерные программы для статистических расчетов «умеют» строить  $Q$ – $Q$  изображение для проверки близости исследуемого эмпирического распределения к нормальному.

Можно сказать, что использование вероятностной бумаги является частным случаем построения  $Q$ – $Q$  изображений.

Важным недостатком графического способа является отсутствие количественных оценок.

**Использование  $\chi^2$ -распределения (критерий Пирсона).** Если число измерений  $N$  достаточно велико, то для проверки близости эмпирического распределения теоретическому можно использовать критерий Пирсона. Заметим, что критерий Пирсона, как и описанный ниже критерий Колмогорова – Смирнова, применим для проверки гипотезы о распределении совокупности при разных (не только нормальных) распределениях. Мы рассмотрим только случай нормально распределенной совокупности. Процедура следующая.

- Для проверки гипотезы о нормальности распределения вначале разбивают диапазон значений измеряемой величины  $X$  на интервалы. Для этого диапазон, в котором лежат значения измеряемой величины  $X$ , делят на  $s$  равных интервалов. Процедура аналогична той, которая выполняется при построении гистограммы распределения (см. выше, разд. 2.2.1). Для повышения надежности разбиения часто присоединяют интервалы, лежащие на краях распределения, к следующим по направлению к центру распределения, если групповая частота, то есть число результатов в интервале, для крайнего интервала меньше 5.

- Подсчитывают число результатов в каждом  $j$ -м интервале  $n_j$  («групповые частоты»). Фиксируют также границы между интервалами  $X'_j$ .
- Для тех же интервалов рассчитывают теоретически ожидаемые групповые частоты  $n_{j,\text{теор}}$ , исходя из предполагаемого распределения генеральной совокупности. В случае, если предполагается нормальное распределение, это делают следующим образом:
  - а) определяют квантили нормированного распределения Гаусса для каждой границы между интервалами (классами)  $x'_j$  по формуле

$$Q_j = \frac{(X'_j - \bar{X})}{s}, \quad (2.29)$$

где  $s$  — стандартное отклонение величин  $X_i$ .

- б) в таблицах нормированного распределения Гаусса находят соответствующие значения функции распределения  $\Phi(Q_j)$ ;
  - в) рассчитывают относительную частоту ( $p_j$ ), ожидаемую для каждого  $j$ -го интервала, как разность  $\Phi(Q_j)$  между верхней  $\Phi(Q_{j,n})$  и нижней  $\Phi(Q_{j,n})$  границами каждого интервала:  $p_j = \Phi(Q_{j,n}) - \Phi(Q_{j,n})$ ;
  - г) находят теоретически ожидаемое число результатов для каждого интервала (класса)  $n_{j,\text{теор}} = Np_j^{25}$ .
- Рассчитывают отклонение  $T$ , характеризующее расхождение между теоретически ожидаемым и эмпирическим распределениями

$$T = \sum_{j=1}^c \left( \frac{n_j - n_{j,\text{теор}}}{n_{j,\text{теор}}} \right)^2. \quad (2.30)$$

Статистическая функция  $T$  подчиняется  $\chi^2$ -распределению с  $f = c - 3$  степенями свободы<sup>26</sup>.

- Полученное значение  $T$  сравнивают с процентной точкой  $\chi^2$ -распределения  $\chi^2(\bar{P}, f)$ , соответствующей доверительной вероятности  $\bar{P}$  и числу степеней свободы  $f = c - 3$  (см. Приложение 3). При  $T > \chi^2(\bar{P}, f)$  гипотезу  $H_0$  отвергают. Критерий Пирсона эффективен при достаточно больших выборках ( $N > 40-50$ ).

**Критерий Колмогорова–Смирнова** используется для проверки той же гипотезы, что и критерий Пирсона, но в отличие от способа Пирсона здесь фактически работают с эмпирической функцией распределения (см. разд. 2.2), оценивая ее отклонение от выбранной теоретической функции распределения.

В случае проверки гипотезы о нормальном распределении совокупности процедура следующая.

Значения  $X_i$  нормируют по формуле  $u_i = (X_i - \bar{X})/s$  и находят значения гауссова интеграла ошибок  $\Phi(u_i)$ , соответствующие  $u_i$ . Рассчитывают разности  $d(X_i) = |F_n(X_i) - \Phi(u_i)|$ , находят максимальную из них  $d_{\max}(X_i)$  и сравнивают  $d_{\max}(X_i)$  с табличным значением  $d(\bar{P}, N)$  для выбранной доверительной вероятности  $\bar{P}$ . Эти ве-

<sup>25</sup> Если предполагается другое распределение совокупности, действуют аналогично с использованием соответствующих таблиц.

<sup>26</sup> В общем случае  $f = c - k$ , где  $k$  — число параметров, необходимых для описания выборки (см. разд. 2.2.2). Для выборок из нормального распределения  $k = 3$  и, соответственно,  $f = c - 3$ .



личины табулированы, например, в [23]. Если  $d_{\max}(X_j) > d(\bar{P}, N)$ , то гипотезу о том, что совокупность имеет нормальное распределение, отвергают.

Критерий Колмогорова – Смирнова относится к непараметрическим (т. е. при его использовании не вводится априорное предположение о распределении результатов измерений) и работает лучше критерия Пирсона при небольших выборках ( $N < 40$ ).

Оба эти способа проверки формы распределения реализованы в большинстве компьютерных программ, предназначенных для статистических расчетов.

**Критерий количества серий.** К методам проверки гипотезы о распределении генеральной совокупности примыкает еще один (непараметрический) метод, используемый для сравнения формы двух распределений при относительно небольшом объеме выборок. Это «критерий количества серий», иногда называемый критерием Вальда – Вольфовица [24].

Предположим, что мы имеем две выборки:

$X_1, X_2, \dots, X_m$  с числом измерений  $m$  и  $Y_1, Y_2, \dots, Y_n$  с числом измерений  $n$ , и хотим проверить гипотезу о том, что обе выборки получены из одной и той же совокупности (без каких-либо предположений о форме ее распределения). Для этого расположим результаты измерений, входящие в обе сравниваемые выборки, в один ряд по возрастанию, после чего обозначим все элементы первой выборки как  $X$ , а второй — как  $Y$ . Будем называть здесь «серией» любую последовательность результатов измерений в общем ряду, принадлежащих к одной выборке.

Обозначим число серий как  $r$ . Критерий количества серий состоит в том, что нулевая гипотеза о принадлежности сравниваемых выборок к одной совокупности должна отвергаться, если число серий  $r$  достаточно мало. Действительно, чем меньше  $r$ , тем хуже «перемешаны» сравниваемые выборки (а значит, имеют различные свойства, принадлежат к совокупностям с разными распределениями).

Для проверки нулевой гипотезы сравнивают найденное число серий  $r$  с критическим значением  $r(P, m, n)$ , зависящим от уровня значимости  $P$  (вероятность ошибки первого рода) и числа измерений в выборках  $m, n$ . Таблица с величинами  $r(P, m, n)$  имеется, например, в [23]. Гипотезу об одинаковой форме распределений совокупностей отвергают при  $r < r(P, m, n)$ .

Отметим, что в большинстве случаев используют двухсторонний критерий (см. разд. 2.1.4), что необходимо учитывать при выборе уровня значимости.

**Пример.** Пусть мы сравниваем форму распределения двух совокупностей, из которых взяты, соответственно, выборка  $X$  ( $m = 9$ ) и выборка  $Y$  ( $n = 10$ ). В упорядоченном ряду элементы выборок выглядят как

$$\frac{XX \ YYY \ X \ YY \ X \ YYYYY \ XX \ Y \ XXX}{1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9},$$

то есть число серий  $r = 9$  (внизу — порядковый номер каждой серии). В [23], табл. 6.7 находим  $r(P = 0,025, 9, 10) = 5$ . Поскольку  $9 > 5$ , гипотезу об одинаковой форме распределений не отвергают с принятой вероятностью ошибки первого рода 0,05.

О критериях сравнения распределений, соответствующих двум выборкам («абсолютная однородность независимых выборок») см. [25].

### 2.4.2. Подозрительно выделяющиеся значения и грубые промахи

Очень часто при рассмотрении результатов многократных измерений какой-либо величины вида (2.5) оказывается, что один или несколько крайних (наибольших или наименьших) результатов измерений существенно отличаются от основной массы результатов («*подозрительно выделяющиеся значения*»). Причиной появления таких результатов может быть грубая погрешность при выполнении конкретного измерения (*грубый промах* или *выброс*, см. разд. 1.3.5). Включение грубых промахов в массив данных для дальнейшей статистической обработки неправомерно и может привести к неприемлемым ошибкам в получаемых оценках и к неправильным выводам при проверке гипотез. Практически, если при выполнении этих измерений не было зафиксировано явных причин промаха (типа «образец уронили на пол»), всегда встает вопрос о том, является ли подозрительно выделяющееся значение грубым промахом или, другими словами, принадлежит ли оно к совокупности, из которой сделана выборка.

Для ответа на этот вопрос прежде всего используют графические способы изображения выборки, в частности строят описанную в разд. 2.2 гистограмму. Очень часто грубые промахи выявляются при этом глазом настолько однозначно, что дальнейшая проверка не требуется. В более сложных случаях проверку гипотезы о принадлежности подозрительно выделяющегося значения к совокупности, из которой получены остальные варианты, проводят с использованием статистических тестов. В большинстве случаев это делается в предположении о нормальной совокупности. Во всех случаях рассматривается упорядоченная выборка.

**Правило грубой ошибки** [26] предполагает нормальное распределение совокупности. Предположим, что в упорядоченной выборке вида (2.5)  $X_1 \leq X_2 \leq \dots \leq X_i \leq \dots \leq X_N$  подозрительно выделяется значение  $X_N$ . Рассчитывают среднее значение  $\bar{X}$  и стандартное отклонение  $s$  для этой выборки (вопрос о том, надо ли учитывать подозрительно выделяющееся значение при расчете  $\bar{X}$  и  $s$ , рассмотрен ниже) и находят величину  $M = (X_N - \bar{X}) / s$  (если подозрительно выделяется значение  $X_1$ , то  $M = (\bar{X} - X_1) / s$ ). При  $M > 4$  подозрительно выделяющееся значение считают грубым промахом<sup>27</sup>. Если в выборке (2.5) подозрительно выделяется значение  $X_1$ , то  $M = (\bar{X} - X_1) / s$ .

При кажущейся простоте применение правила грубой ошибки во многих ситуациях приводит к неоднозначным выводам.

**Пример.** В результате пяти выполненных в одних условиях измерений концентрации меди в сплаве получены следующие значения (в процентах):

$$42,9; 43,3; 43,5; 43,6; \mathbf{44,8}. \quad (2.31)$$

Предположим, что в данной выборке подозрительно выделяющееся значение 44,8 является выбросом, и не будем включать его в расчет  $\bar{X}$  и  $s$ . Тогда получим:  $\bar{X} = 43,32$ ,  $s = 0,31$  и  $M = 4,77 > 4$ , т. е. результат **44,8** действительно является грубым промахом.

<sup>27</sup> Иногда выброс регистрируют при  $M > 3$ . Кроме того, в [27] приведены критические значения «критерия Граббса» для той же статистики  $|X_{\text{выб}} - \bar{X}| / s$ .



Предположим теперь, что значение **44,8** не является выбросом, и включим его в расчеты. Тогда  $\bar{X} = 43,62$ ,  $s = 0,71$  и  $M = 1,66 < 4$ , т. е. результат **44,8** не является грубым промахом. Что захотим, то и получим!

Таким образом, при использовании правила грубой ошибки может возникать неоднозначность, особенно при малых выборках. Поэтому правило грубой ошибки обычно применяют при достаточно больших выборках ( $N > 20-30$ ), редко встречающихся при химическом анализе. При этом для принятия решения об отбросе того или иного значения  $X_i$  нужно, чтобы результат теста не зависел от того, учитывается ли подозрительно выделяющееся значение при расчете величины  $M$  или нет. Такая проблема не возникает при использовании теста Диксона.

Тест Диксона («Dixon's Q») [28] используют при небольших объемах выборки ( $N < 20$ ) и в предположении о нормальном распределении генеральной совокупности.

Рассчитывают величину  $Q_{i,k}$  по одной из формул, приведенных в табл. 2.2, причем статистическую функцию  $Q_{i,k}$  выбирают в зависимости от общего числа измерений  $N$ .

Найденные значения  $Q_{i,k}$  сравнивают с табличным критическим значением  $Q_{i,k,\text{крит}}$  соответствующим  $N$  и выбранной вероятности ошибки второго рода  $\beta$  – «степени риска» (Приложение 6). Если  $Q_{i,k} > Q_{i,k,\text{крит}}$ , то подозрительно выделяющееся значение считают грубым промахом (выбросом) и не используют при дальнейшей статистической обработке результатов измерений.

Применим тест Диксона к примеру (2.31). Из табл. 2.2 ясно, что при  $N = 5$  надо брать  $Q_{1,0}$ . Находим  $Q_{1,0} = (X_N - X_{N-1}) / (X_N - X_1) = (44,8 - 43,6) / (44,8 - 42,9) = 0,632$ . Это меньше табличного критического значения  $Q_{1,0} = 0,642$  для  $N = 5$  и вероятности ошибки второго рода 0,05. Следовательно, значение 44,8 не является грубым промахом при проверке по тесту Диксона.

Таблица 2.2. Формулы для расчета  $Q_{i,k}$

$N$	$Q_{i,k}$	Выделяющееся значение	
		$X_N$	$X_1$
$3 \leq N \leq 7$	$Q_{1,0}$	$(X_N - X_{N-1}) / (X_N - X_1)$	$(X_2 - X_1) / (X_N - X_1)$
$8 \leq N \leq 10$	$Q_{1,1}$	$(X_N - X_{N-1}) / (X_N - X_2)$	$(X_2 - X_1) / (X_{N-1} - X_1)$
$11 \leq N \leq 13$	$Q_{2,0}$	$(X_N - X_{N-2}) / (X_N - X_2)$	$(X_3 - X_1) / (X_{N-1} - X_1)$
$14 \leq N \leq 25$	$Q_{2,2}$	$(X_N - X_{N-2}) / (X_N - X_3)$	$(X_3 - X_1) / (X_{N-2} - X_1)$

**MAD-тест** [29]. Основная проблема, возникающая при использовании правила грубой ошибки и теста Диксона, связана с возможностью неправильной оценки стандартного отклонения  $s$  в случае, когда гипотеза о нормальном распределении совокупности неверна. Если сведения о распределении совокупности отсутствуют, можно применить MAD-тест, относящийся к робастным, т. е. устойчивым к нарушениям основных предположений, методам. Для оценки дисперсии выборки в этом методе используется медиана абсолютных отклонений median absolute deviation, MAD:

$$\text{MAD} = \text{median}[|X_i - \text{median}(X)|], \quad (2.32)$$

где *median* – обозначение медианы (см. разд. 2.2.2).

Для проверки гипотезы о том, что подозрительно выделяющееся значение  $X_{\text{выд}}$  (это может быть как  $X_N$ , так и  $X_1$ ) является грубым промахом, рассчитывают отношение

$$\frac{|X_{\text{выд}} - \text{median}(X_i)|}{MAD}. \quad (2.33)$$

Если это отношение больше 5, то подозрительно выделяющееся значение  $X_{\text{выд}}$  является грубым промахом (выбросом) и должно быть отброшено.

Рассмотрим применение теста на примере (2.31). Медиана  $\text{median}(X_i)$  равна 43,5, величины  $|X_i - \text{median}(X_i)|$  составляют 0,6; 0,2; 0; 0,1; 1,3. Медиана этих величин  $MAD = 0,2$ , отношение (2.33) равно  $(44,8 - 43,5)/0,2 = 6,5 > 5$ , т. е. результат измерения 44,8 является грубым промахом. Данное заключение противоречит выводу, сделанному на основании теста Диксона, что, очевидно, является следствием отличного от нормального распределения генеральной совокупности.

Существует несколько правил, соблюдение которых необходимо для получения корректных результатов при выявлении грубых промахов:

- как и другие методы математической статистики, описанные выше тесты для выявления грубых промахов применимы только к некоррелированным данным и их нельзя применять к взаимозависимым результатам измерений;
- к каждой выборке может применяться любой подходящий тест, но только один раз. Многократное использование какого-либо теста (например, теста Диксона) некорректно и ведет к грубым ошибкам, во избежание которых необходимо перейти к моделированию по методу Монте-Карло;
- недопустим произвольный отброс подозрительно выделяющихся значений;
- хотя выявленные грубые промахи не учитываются при статистической обработке результатов измерений, однако они не должны забываться вообще
- их показывают на гистограммах, сообщают при выдаче результатов измерений заказчикам и т. д. Кроме того, каждый выброс должен анализироваться с точки зрения причин его появления.

Вышеописанные тесты предназначены для выявления одиночных выбросов в одномерной выборке. Такая же проблема может возникнуть и для двумерной выборки (в частности, при построении градуировочной зависимости). Проблема «выбросов» для двумерной выборки (при регрессионном анализе) рассмотрена в разд. 2.7.2.4.

Кроме того, могут «выпасть» два или более близких значения. Конечно, в практике измерений (особенно при химическом анализе) в силу ограниченности объемов выборок подобные случаи бывают редко. Тем не менее критерии отброса для таких случаев разработаны (см., например, [28]).

Другим, хотя и не универсальным, решением проблемы подозрительно выделяющихся значений является использование при обработке данных робастных методов [30, 31], одним из которых является MAD-тест.

Описанные выше методы выявления грубых промахов применимы лишь при достаточно больших выборках (обычно при числе измерений  $N$  не менее 5). Однако абсолютное большинство методик выполнения измерений предусматривают меньшее число измерений (чаще всего  $N = 2$ ). При таких измерениях для выявле-



ния ситуаций, связанных с наличием грубых промахов, и расчетов окончательно приводимых результатов измерений используют специальные приемы, связанные с использованием медианы и выполнением дополнительных измерений и применимые тогда, когда стандартное отклонение известно («проверка приемлемости», см. разд. 4.4)

## 2.5. Некоторые наиболее часто встречающиеся статистические гипотезы и способы их проверки

Как указывалось в разд. 2.1.4, многие вопросы, касающиеся свойств исследуемых объектов и решаемые по результатам обработки данных измерений, формулируются в форме гипотез. Ниже рассмотрены несколько типичных гипотез и способы их проверки. Эти гипотезы часто проверяются как при решении метрологических задач, так и при обработке результатов измерений в других целях.

### 2.5.1. Гипотеза о равенстве двух дисперсий

Пусть мы имеем две независимые выборки:

$X_1, X_2, \dots, X_{N_1}$  (число измерений  $N_1$ , стандартное отклонение  $s_x$ ) и  $Y_1, Y_2, \dots, Y_{N_2}$  (число измерений  $N_2$ , стандартное отклонение  $s_y$ ).

Предположим, что первая выборка получена из *нормальной* совокупности с дисперсией, вторая — из *нормальной же* совокупности с дисперсией  $\sigma_y^2$ . Необходимо проверить гипотезу  $H_0: (\sigma_x^2 = \sigma_y^2 = \sigma^2)$ , т. е. гипотезу о том, что обе выборки получены из нормальных совокупностей с одинаковой дисперсией  $\sigma^2$ . Проверку гипотезы проводят с использованием *F-критерия Фишера*.

Для случая, когда обе выборки получены из нормальных совокупностей, величина  $F = s_x^2/s_y^2$  (где  $s_x \geq s_y$ ) подчиняется *F-распределению* с числами степеней свободы  $f_1 = N_1 - 1, f_2 = N_2 - 1$ . Гипотеза  $H_0: (\sigma_x^2 = \sigma_y^2 = \sigma^2)$  отвергается с принятой доверительной вероятностью  $\bar{P}$  при  $F > F(\bar{P}, f_1, f_2)$ , где  $F(\bar{P}, f_1, f_2)$  — процентная точка *F-распределения*, т. е. предел интегрирования -распределения при принятой доверительной вероятности  $\bar{P}$  и числах степеней свободы  $f_1$  и  $f_2$  (Приложение 5).

Многие современные компьютерные программы непосредственно рассчитывают величину доверительной вероятности или величину вероятности ошибки 1-го рода  $\alpha = (1 - \bar{P})$ , часто обозначаемую просто как  $P$ . Тогда гипотезу  $H_0$  отвергают, если  $P < P_{\text{крит}}$ , где  $P_{\text{крит}}$  — заранее выбранное значение вероятности ошибки 1-го рода.

Имея выборки не слишком большого размера, мы можем лишь предполагать нормальность распределения генеральной совокупности. Критерии, при использовании которых предполагается некоторое точно известное (чаще всего нормальное) распределение генеральной совокупности, называют *параметрическими*. *F-критерий* относится к параметрическим критериям, так как для его применения необходимо быть уверенным в нормальном распределении совокупностей.

На результаты проверки по *F-критерию* сильно влияют грубые промахи — наличие даже одного выброса в сравниваемых результатах измерений часто приводит к неправильным выводам при проверке гипотезы  $H_0$ . Это связано с тем, что при



расчете стандартного отклонения суммируются *квадраты* разностей измеренных величин и среднего значения. Поэтому прежде чем применять  $F$ -критерий, обе сравниваемые выборки необходимо исследовать на наличие подозрительно выделяющихся значений (см. разд. 2.4.2) и, если они окажутся выбросами, исключить их из рассмотрения.

Различие между сравниваемыми дисперсиями, достоверно выявляемое с помощью  $F$ -критерия при одной и той же доверительной вероятности, уменьшается с ростом числа измерений  $N_1$  и  $N_2$ . Если объемы выборок (число измерений для каждой выборки) можно планировать, то лучше предусмотреть большее число измерений для выборки, имеющей меньшее стандартное отклонение.

Гипотезу о равенстве двух дисперсий проверяют, в частности, в ходе проверки гипотезы о равенстве двух математических ожиданий при несвязанных выборках (см. разд. 2.5.3.1 и соответствующий пример). Сравнение двух дисперсий по критерию Фишера лежит также в основе дисперсионного анализа (см. ниже, разд. 2.6 и сл.).

### 2.5.2. Сравнение нескольких дисперсий

Пусть мы имеем  $m$  ( $m > 3$ ) независимых выборок, и каждая выборка содержит не менее двух ( $N > 2$ ) результатов измерений:

Таблица 2.3. Результаты исследования нескольких выборок

Номер выборки	Результаты измерений, соответствующие номерам измерений						Стандартное отклонение выборки
	1	2	...	$i$	...	$N_j$	
1	$X_{11}$	$X_{12}$	...	$X_{1i}$	...	$X_{1N_1}$	$s_1$
2	$X_{21}$	$X_{22}$	...	$X_{2i}$	...	$X_{2N_2}$	$s_2$
...	...	...	...	...	...	...	...
$J$	$X_{J1}$	$X_{J2}$	...	$X_{Ji}$	...	$X_{JN_J}$	$s_J$
...	...	...	...	...	...	...	...
$M$	$X_{m1}$	$X_{m2}$	...	$X_{mi}$	...	$X_{mN_m}$	$s_m$

Обозначим стандартные отклонения этих выборок как  $s_1, s_2, \dots, s_m$ . Предположим также, что все эти выборки получены из нормальных совокупностей со стандартными отклонениями  $\sigma_1, \sigma_2, \dots, \sigma_m$ . Необходимо проверить гипотезу  $H_0$ : ( $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2 = \sigma^2$ ), т. е. гипотезу о том, что все рассматриваемые выборки получены из нормальных совокупностей с одинаковой дисперсией  $\sigma^2$ . Проверки гипотезы может быть проведена с использованием *критерия Бартлетта, или М-критерия*. Процедура проверки состоит в следующем:

- рассчитывают величину

$$\begin{aligned}
 S &= \sqrt{\frac{\sum_{i=1}^{N_1} (X_{1i} - \bar{X}_1)^2 + \sum_{i=1}^{N_2} (X_{2i} - \bar{X}_2)^2 + \dots + \sum_{i=1}^{N_m} (X_{mi} - \bar{X}_m)^2}{N - m}} = \\
 &= \sqrt{\frac{\sum_{j=1}^m \sum_{i=1}^{N_j} (X_{ji} - \bar{X}_j)^2}{N - m}},
 \end{aligned} \tag{2.34}$$

где  $N = \sum_{j=1}^m N_j$  — общее число измерений во всех выборках,  $(N - m) = f_g$  — общее число степеней свободы, фактически равное сумме числа степеней свободы всех выборок  $f_g = \sum_{j=1}^m f_j \sum_{j=1}^m (N_j - 1)$ .

— находят величину

$$M^* = 2,303 \left( f_g \lg S^2 - \sum_{j=1}^m \lg s_j^2 \right). \quad (2.35)$$

С достаточной для большинства практических целей точностью эта величина характеризуется  $\chi^2$ -распределением с числом степеней свободы  $f = m - 1$ . Для проверки гипотезы  $H_0: (\sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2 = \sigma^2)$  сравнивают найденную величину  $M^*$  с процентной точкой  $\chi^2$ -распределения  $\chi^2(\bar{P}, f)$ , соответствующей доверительной вероятности  $\bar{P}$  и числу степеней свободы  $f = m - 1$ . При  $M^* > \chi^2(\bar{P}, f)$  гипотезу  $H_0$  отвергают.

Более точные результаты получаются при использовании вместо  $M^*$  величины  $M = M^*/C$ , то есть после введения поправки  $C$ , равной

$$C = \frac{\left[ \sum_{j=1}^m \left( \frac{1}{f_j} \right) - \frac{1}{f_g} \right]}{3(m-1)} + 1. \quad (2.36)$$

Эту поправку вносят при расчетах, проводимых с помощью калькулятора, в случаях, когда рассчитанное значение  $M^*$  не намного превышает  $\chi^2(\bar{P}, f)$ . Современные компьютерные программы всегда рассчитывают поправленную величину  $M$ .

Еще раз заметим, что многие компьютерные программы выводят непосредственно значение вероятности ошибки первого рода  $\alpha$ , часто обозначаемой как  $P$ . В этом случае гипотезу  $H_0$  отвергают, если  $P$  меньше принятой вероятности ошибки первого рода  $P = 1 - \bar{P}$ .

**Критерий Кохрена (Коцрена, Cochren).** Пусть мы имеем ту же совокупность результатов измерений, к которой применялась проверка по критерию *Бартлетта* (см. табл. 2.3), но число результатов измерений во всех выборках одно и то же ( $N_1 = N_2 = \dots = N_m = N$  и, соответственно,  $f_1 = f_2 = \dots = f_m = f$ ). Упорядочим выборки по убыванию их стандартного отклонения  $s_j: s_m \geq s_{m-1} \geq \dots \geq s_1$ . Предположим, что все выборки, кроме выборки со стандартным отклонением  $s_m$  (то есть с наибольшей дисперсией  $s_m^2$ ), получены из нормально распределенных совокупностей с одинаковыми дисперсиями  $\sigma^2$  ( $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_{m-1}^2 = \sigma^2$ ), а выборка с дисперсией  $s_m^2$  — также из нормально распределенной совокупности с дисперсией  $\sigma_m^2$ . Проверим гипотезу  $H_0: (\sigma_m^2 = \sigma^2)$ , т. е. гипотезу о том, что выборка с максимальной оценкой дисперсии (наибольшим стандартным отклонением) получена из нормальной совокупности с той же дисперсией  $\sigma^2$ , что и остальные выборки. В качестве альтернативной используется гипотеза  $H_1: (\sigma_m^2 > \sigma^2)$ . Проверка проводится следующим образом.

Рассчитывают величину

$$G = \frac{s_m^2}{s_1^2 + s_2^2 + \dots + s_m^2}. \quad (2.37)$$

Найденную величину  $G$  сравнивают с процентной точкой распределения  $g(P, m, f = N - 1)$ , где  $P$  – принятая вероятность ошибки 1-го рода,  $m, f$  – числа степеней свободы. Такие таблицы имеются, в частности, в [23]. При  $G > g(P, m, f = N - 1)$  гипотезу  $H_0: (\sigma_m^2 = \sigma^2)$  отвергают.

Критерий Кохрена используют как для проверки однородности дисперсий, так и для выявления грубых ошибок. В последнем случае выборку (серию измерений) с наибольшим стандартным отклонением фактически рассматривают как подозрительно выделяющийся результат. Если при проверке по критерию Кохрена гипотеза  $H_0: (\sigma_m^2 = \sigma^2)$  отвергается, эту выборку отбрасывают. Такой прием используют, в частности, при оценке качества химического анализа (см. разд. 4.5.1.2).

Отметим, что критерии сравнения нескольких дисперсий, особенно критерий Бартлетта, очень чувствительны как к наличию грубых промахов, так и к нарушению предположения о нормальном распределении генеральных совокупностей. Если часть совокупностей, из которых получены выборки, имеет отличные от нормального распределения, то гипотеза о равенстве дисперсий при проверке по критерию Бартлетта может быть отвергнута с *большой* вероятностью. То же касается и использования критерия Кохрена. Поэтому к результатам использования этих критериев надо относиться с осторожностью, особенно при небольших по объему выборках.

### 2.5.3. Гипотеза о равенстве двух математических ожиданий при независимых (несвязанных) выборках

Это одна из наиболее часто встречающихся гипотез.

Имеем две независимые выборки:

$X_1, X_2, \dots, X_r, \dots, X_{N_1}$  (число измерений  $N_1$ , среднее значение  $\bar{X}$ , стандартное отклонение  $s_X$ ) и

$Y_1, Y_2, \dots, Y_r, \dots, Y_{N_2}$  (число измерений  $N_2$ , среднее значение  $\bar{Y}$ , стандартное отклонение  $s_Y$ ).

Предположим, что эти выборки получены из совокупностей с математическими ожиданиями  $\mu_X$  и  $\mu_Y$ . Тогда проверяемая гипотеза  $H_0: (\mu_X = \mu_Y = \mu)$ .

Каждая из рассматриваемых здесь выборок представляет собой результаты полученных в одинаковых условиях результатов измерений<sup>28</sup>. С точки зрения математической статистики эта же задача решается при:

- проверке гипотезы о равенстве математических ожиданий результатов измерений концентрации какого-либо вещества в одном образце, выполненных в двух лабораториях либо двумя лаборантами;
- проверке гипотезы об отсутствии возрастных различий (в среднем) между студентами двух групп

и так далее.

**Пример.** Имеются две колбы с водными растворами хлорид-иона. Необходимо проверить предположение о том, что концентрация его в обеих колбах одинакова. Для проверки

<sup>28</sup> Заметим, что такую задачу часто называют «сравнением двух средних», что не вполне корректно – гипотезы проверяются не относительно экспериментально найденных средних значений (которые точно известны), а относительно параметров генеральных совокупностей, отражающих свойства исследуемых объектов (см. выше, разд. 2.1.4).



проводится несколько измерений концентрации в каждой колбе объемным argentометрическим методом.

### 2.5.3.1. Случай нормального распределения совокупностей

В предположении о нормальном распределении совокупностей, соответствующих каждой из выборок, задача решается с использованием *t-критерия (критерия Стьюдента)*.

Полагаем, что выборки получены из следующих нормальных совокупностей: первая — с математическим ожиданием  $\mu_x$  и стандартным отклонением  $\sigma_x$ , вторая — с математическим ожиданием  $\mu_y$  и стандартным отклонением  $\sigma_y$ . Это означает, что результаты измерений, входящие в каждую выборку, описываются моделью (2.4). Необходимо проверить гипотезу  $H_0: (\mu_x = \mu_y = \mu)$ , т. е. гипотезу о том, что разница между средними значениями измеряемой величины  $\bar{X}$  и  $\bar{Y}$  есть лишь следствие случайных погрешностей. Процедуру проверки гипотезы проводят в следующей последовательности:

- Проверяют гипотезу о том, что обе выборки получены из нормальных совокупностей с одинаковой дисперсией  $H_0: (\sigma_x^2 = \sigma_y^2 = \sigma^2)$ . Способ проверки с использованием критерия Фишера описан выше в разд. 2.5.1.

Если гипотеза о равенстве дисперсий не отвергается, рассчитывают стандартное отклонение разности средних значений  $(\bar{X} - \bar{Y})$  сравниваемых выборок:

$$s = \sqrt{\frac{s_x^2(N_1 - 1) + s_y^2(N_2 - 1)}{N_1 + N_2 - 1}} = \sqrt{\frac{s_x^2 f_1 + s_y^2 f_2}{f_1 + f_2}}, \quad (2.38)$$

где  $f_1 = N_1 - 1$  и  $f_2 = N_2 - 1$  — числа степеней свободы для рассматриваемых выборок.

- Рассчитывают величину

$$N_2 = \frac{|\bar{X} - \bar{Y}|}{s} \sqrt{\frac{N_1 N_2}{N_1 + N_2}}, \quad (2.39)$$

которая имеет *t*-распределение с  $f = N_1 + N_2 - 1$  степенями свободы.

- Вычисленную величину *t* сравнивают с процентной точкой *t* — распределения Стьюдента  $t(\bar{P}, f)$ , (см. Приложение 4), соответствующей выбранной доверительной вероятности  $\bar{P}$  и числу степеней свободы  $f = N_1 + N_2 - 1$ . Если  $t > t(\bar{P}, f)$ , то гипотезу  $H_0: (\mu_x = \mu_y = \mu)$  отвергают.

Особо следует отметить, что доверительную вероятность выбирают, учитывая, какова альтернативная гипотеза (см. разд. 2.1.4).

Если априорная информация о возможном соотношении  $\mu_x$  и  $\mu_y$  отсутствует, проводят проверку по *двухстороннему критерию*. Это соответствует альтернативной гипотезе  $H_1: (\mu_x \neq \mu_y)$ . В этом случае выбирают ошибку первого рода  $\alpha = (1 - \bar{P})$  в два раза меньшую, чем заранее выбранная вероятность ошибки 1-го рода.

Если же заранее известно, что какая-либо из величин  $\mu_x$  и  $\mu_y$  больше другой, применяют *односторонний критерий*. В этом случае альтернативная гипотеза  $H_1: (\mu_x > \mu_y)$ . При расчетах используют ту вероятность ошибки первого рода  $\alpha$ , которая была выбрана для проверки основной гипотезы. Так, если в нашем примере заранее

известно, что концентрация хлорид-иона в первой колбе не меньше, чем во второй, то при расчетах мы должны использовать заранее выбранную вероятность  $\alpha = 0,05$  (доверительную вероятность  $\bar{P} = 0,95$ ).

Другими словами, если при проверке гипотезы  $H_0: (\mu_x = \mu_y = \mu)$  выбрана доверительная вероятность  $\bar{P}$ , то при использовании одностороннего критерия берут процентную точку  $t$ -распределения Стьюдента  $t(\bar{P}, f)$ , а двухстороннего –  $t\left[1 - \frac{1 - \bar{P}}{2}, f\right]$ .

Так, если в нашем примере о концентрации хлорид-иона в колбах ничего заранее не известно и мы выбрали вероятность ошибки 1-го рода  $\alpha = 0,05$  (доверительную вероятность 0,95), то при расчетах мы должны использовать  $\alpha = 0,025$  (доверительную вероятность 0,975).

Если гипотеза о равенстве дисперсий была отвергнута, значение  $t$  рассчитывают не по (2.36), а по формуле

$$t = \frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{s_x^2}{N_1} + \frac{s_y^2}{N_2}}}. \quad (2.40)$$

Величина  $t$ , рассчитанная по (2.40), имеет распределение, близкое к распределению Стьюдента с числом степеней свободы

$$f = \frac{\left(\frac{s_x^2}{N_1} + \frac{s_y^2}{N_2}\right)^2}{\frac{\left(\frac{s_x^2}{N_1}\right)^2}{N_1 - 1} + \frac{\left(\frac{s_y^2}{N_2}\right)^2}{N_2 - 1}}. \quad (2.41)$$

Вычисленную по (2.41) величину  $f$  округляют до ближайшего целого числа.

Собственно проверку гипотезы  $H_0: (\mu_x = \mu_y = \mu)$  проводят так же, как в случае, когда гипотеза о равенстве дисперсий не отвергается, с тем отличием, что величину  $t$  находят по формуле (2.40), а процентную точку  $t$ -распределения Стьюдента  $t(\bar{P}, f)$  – для числа степеней свободы  $f$ , рассчитанного по формуле (2.41).

**Пример.** Имеются две колбы с водными растворами хлорид-иона. Необходимо было проверить гипотезу о том, что концентрация его в обеих колбах одинакова, для чего было выполнено несколько измерений концентрации в каждой колбе объемным argentометрическим методом. Измерения в разных колбах выполнялись разными лаборантами. Известно, что результаты измерений по данному методу имеют близкое к нормальному распределение. Получены следующие результаты (в ммоль/л):

$$\left. \begin{array}{l} \text{Колба 1: } 5,12; 5,16; 4,05; 5,03; 5,09 \\ \text{Колба 2: } 5,01; 5,04; 5,06; 5,03; 5,03 \end{array} \right\} \quad (2.42)$$

Результат 4,05 является грубым промахом по любому из тестов разд. 2.4.2. Кроме того, лаборант зафиксировал, что при этом измерении он мог ошибиться. Поэтому результат 4,05 из дальнейших расчетов исключаем. Для решения задачи необходимо проверить гипотезу  $H_0: (\mu_x = \mu_y = \mu)$ , где  $\mu_x$  и  $\mu_y$  – математические ожидания выборок для двух колб.

Средние значения и стандартные отклонения выборок равны:

$$\left. \begin{array}{l} \text{Колба 1. } \bar{X} = 5,100, s_X = 0,0548 (N_1 = 4) \\ \text{Колба 2. } \bar{Y} = 5,034, s_Y = 0,0182 (N_2 = 5) \end{array} \right\} \quad (2.43)$$

Вначале проверяем гипотезу о равенстве двух дисперсий (примем доверительную вероятность  $\bar{P} = 0,95$ ). Находим  $F = s_X^2/s_Y^2 = 0,09 > 5,19$ , где 5,19 есть  $F(0,95, 4, 5)$  (см. Приложение 3). Гипотеза о равенстве двух дисперсий отвергается (что, вероятно, связано с тем, что измерения выполнялись разными лаборантами), и при расчетах необходимо использовать формулы (2.40, 2.41).

**Вариант 1.** Дополнительной информации нет, и в качестве альтернативной используется гипотеза  $H_1: (\mu_X \neq \mu_Y)$ . Находим  $t = 2,310, f = 3,533 \approx 4$ . Критическое значение  $t(0,975, 4) = 2,776 > 2,310$ . Таким образом, гипотеза о равенстве концентраций иона натрия в колбах не отвергается с принятой доверительной вероятностью  $\bar{P} = 0,95$  (двухсторонний критерий).

**Вариант 2.** Имеются сведения, что исходно в колбах находился один и тот же раствор, и перед измерением в первую колбу уронили маленький кристаллик неизвестной соли, объемом которого можно пренебречь. Тогда концентрация натрия в первой колбе не может быть меньше концентрации во второй, и в качестве альтернативной необходимо использовать гипотезу  $H_1: (\mu_X > \mu_Y)$ . Величины  $t = 2,310$  и  $f = 3,533 \approx 4$  те же. Критическое значение  $t(0,97, 4) = 2,132 < 2,310$ , то есть гипотеза о равенстве математических ожиданий отвергается с  $\bar{P} = 0,95$ . Различие концентраций иона натрия в двух колбах выявлено достоверно с  $\bar{P} = 0,95$  (односторонний критерий). Причиной различия в выводах с вариантом 1 — использование дополнительной информации.

Некоторые, не слишком значительные отклонения распределений совокупностей от нормального мало влияют на результаты проверки гипотезы  $H_0: (\mu_X = \mu_Y = \mu)$  с использованием критерия Стьюдента. Большую опасность представляют грубые промахи. Поэтому при наличии в сравниваемых выборках подозрительно выделяющихся значений необходимо проверять, являются ли эти значения выбросами и если да, удалять их из расчетов.

Проверка гипотезы о равенстве математического ожидания выборки точно известной величине. Рассмотрим ситуацию, когда имеется только одна выборка  $X_1, X_2, \dots, X_N$  (число измерений  $N$ , среднее значение  $\bar{X}$ , стандартное отклонение  $s_X$ ), полученная из нормальной совокупности с математическим ожиданием  $\mu_X$ , и необходимо проверить гипотезу  $H_0: (\mu_X = \mu)$ , где  $\mu$  — точно известная величина. Такая ситуация имеет место, например, при анализе стандартного образца (см. разд. 3.2), погрешность аттестации которого пренебрежимо мала. Фактически это — частный случай рассмотренной выше задачи. Решение можно найти из формулы (2.39), в которой надо устремить  $N_2$  к бесконечности и принять  $s_Y = 0$ . Получаем формулу

$$t = \frac{|\bar{X} - \mu|}{s_X} \sqrt{N}. \quad (2.44)$$

Величина  $t$ , рассчитанная по (2.42), имеет распределение Стьюдента с числом степеней свободы  $f = N - 1$ ; проверка гипотезы  $H_0: (\mu_X = \mu)$  проводится аналогично проверке гипотезы о равенстве двух математических ожиданий (см. выше).

**Пример.** Для проверки правильности методики определения кадмия методом атомно-эмиссионного анализа с индукционно-связанной плазмой (ИСП-АЭ) использовали государственный стандартный образец ГСО 6690-93. Это — водный раствор  $\text{Cd}(\text{NO}_3)_2$  с концен-



трацией ионов кадмия ( $1,000 \pm 0,010$ ) г/дм<sup>3</sup>. ГСО разбавили в 200 раз бидистиллированной водой с достаточно высокой точностью, получив водный раствор с концентрацией ионов кадмия ( $5,00 \pm 0,05$ ). Погрешностью концентрации полученного раствора ( $\pm 0,05$  мг/дм<sup>3</sup>) по сравнению с погрешностью метода ИСП-АЭ (не менее 5; отн.) пренебрегаем, полагая концентрацию равной точно 5,00 мг/дм<sup>3</sup>.

Выполнили 7 измерений, получив следующие результаты  $X_i$ :

5,11; 5,20; 4,96; 4,99; 5,08; 5,38; 5,14 (мг/дм<sup>3</sup>). Для проверки правильности необходимо проверить гипотезу  $H_0$ : ( $\mu_x = 5,00$ ). Подозрительно выделяющееся значение 5,38 не является грубым промахом, поэтому считаем полученную выборку однородной. Находим:  $\bar{X} = 5,1228$ ,  $s_x = 0,1406$ ,  $t = 2,311$ . Табличное значение  $t(0,975, 6) = 2,571$  (см. Приложение 4), что соответствует вероятности ошибки 1-го рода 0,05 для двухстороннего критерия. Экспериментально полученное значение 2,311 меньше табличного 2,571, то есть гипотеза  $H_0$  не отвергается — мы не обнаружили систематической погрешности метода с принятой доверительной вероятностью.

При проверке гипотезы о равенстве двух математических ожиданий при несвязанных выборках по  $t$ -критерию Стьюдента выше был использован первый из двух упомянутых в разд. 2.1.4 способов проверки гипотез, а именно статистический тест. Для решения той же задачи можно использовать и второй способ — с расчетом доверительных интервалов. Для этого необходимо рассчитать границы этих интервалов.

Пусть мы имеем выборку вида (2.1), то есть совокупность из  $N$  результатов измерений какой-либо величины. Если справедлива модель (2.4), то оценкой математического ожидания измеряемой величины  $X$  является среднее значение  $\bar{X}$ . Доверительным интервалом для величины  $X$  называют интервал, в котором это математическое ожидание (в отсутствие систематических погрешностей оно же истинное значение измеряемой величины) находится с доверительной вероятностью  $\bar{P}$ . С другой стороны, если бы мы имели бесконечно большое число выборок того же объема из той же совокупности (провели бы такие же измерения в тех же условиях множество раз), то, например, при  $\bar{P} = 0,95$  среднее значение оказалось бы внутри доверительного интервала в 95 случаях из 100.

В случае нормальной совокупности (2.1) со средним значением  $\bar{X}$  и стандартным отклонением  $s$  нижняя  $X_n$  и верхняя  $X_v$  границы доверительного интервала рассчитываются по формулам

$$X_n = \bar{X} - \frac{s \cdot t(\bar{P}, f)}{\sqrt{N}}, \quad X_v = \bar{X} + \frac{s \cdot t(\bar{P}, f)}{\sqrt{N}}, \quad (2.45)$$

где  $t(\bar{P}, f)$  — процентная точка  $t$  — распределения Стьюдента, соответствующая выбранной доверительной вероятности  $\bar{P}$  и числу степеней свободы  $f = N - 1$ . Таким образом, с доверительной вероятностью  $\bar{P}$  измеряемая величина лежит в интервале  $(X_n, X_v)$ .

Очень часто доверительный интервал записывают как  $(\bar{X} \pm \Delta X)$ ; для нормально распределенной величины это выглядит как

$$\left( \bar{X} \pm \frac{s \cdot t(\bar{P}, f)}{\sqrt{N}} \right). \quad (2.46)$$

В этом случае надо иметь в виду следующее. Такая форма записи создает иллюзию, что математическое ожидание (истинное значение измеряемой величины)

лежит скорее ближе к середине доверительного интервала (к величине  $\bar{X}$ ), чем к его краям. Но это не так — можно утверждать только, что истинное значение измеряемой величины находится где-то в пределах доверительного интервала, и не более того. При несимметричных распределениях (например, логарифмически нормальном распределении) границы доверительного интервала расположены несимметрично относительно среднего значения и число, расположенное в середине интервала, не имеет физического смысла.

Способы расчета границ доверительного интервала для распределений, отличающихся от нормального, даны, например, в [26].

При проверке гипотез о равенстве двух математических ожиданий с использованием доверительных интервалов рассчитывают доверительные интервалы для обеих измеряемых величин ( $X_n, X_n$ ) и ( $Y_n, Y_n$ ) при выбранном уровне доверительной вероятности. Если эти интервалы не перекрываются, гипотеза  $H_0: (\mu_X = \mu_Y = \mu)$  отвергается, и наоборот.

Гипотезу о равенстве математического ожидания генеральной совокупности точно известной величине также можно проверить по доверительному интервалу. Для этого по результатам измерений рассчитывают доверительный интервал ( $X_n, X_n$ ). Гипотеза  $H_0: (\mu_X = \mu)$  отвергается (при выбранном уровне доверительной вероятности), если точно известное математическое ожидание  $\mu$  не лежит в этом интервале.

По доверительным интервалам иногда осуществляют проверку других гипотез, например, гипотезы о равенстве двух дисперсий. При этом применяют специальные способы расчета границ доверительных интервалов.

### 2.5.3.2. Непараметрический критерий Вилкоксона — Манна — Уитни

Рассмотренные выше способы проверки гипотезы о равенстве математических ожиданий относятся к параметрическим, то есть к таким, для которых предполагается известной функция распределения совокупностей и, соответственно, характеризующие ее параметры. Так, использование  $t$ -критерия Стьюдента предполагает нормальное распределение величин  $X$  и  $Y$ :  $X \in N(\mu_X, \sigma_X^2)$  и  $Y \in N(\mu_Y, \sigma_Y^2)$ . В свою очередь эти распределения полностью описываются параметрами  $\mu_X, \sigma_X$  и  $\mu_Y, \sigma_Y$ .

К сожалению, вид распределений зачастую неизвестен и, соответственно, никакими параметрами распределений оперировать нельзя. В таких случаях для решения задач математической статистики можно использовать *непараметрические методы*. Поскольку эти методы не предполагают каких-либо параметров распределений, их называют также свободными от параметров или *свободно распределенными*<sup>29</sup>.

Для проверки гипотезы о равенстве двух математических ожиданий при независимых (несвязанных) выборках используется *непараметрический критерий U* («критерий Вилкоксона — Манна — Уитни»). Его проще всего описать на примере.

**Пример.** В двух лабораториях проведены многократные измерения концентрации серы в стандартном образце авиационного смазочного масла. В результате получены две независимые выборки (результаты выражены в массовых процентах):

<sup>29</sup> Использование непараметрических методов имеет, тем не менее, некоторые ограничения. Так, большинство таких методов основано на предположении об унимодальности распределения каждой совокупности (то есть о том, что плотность распределения имеет один максимум).



Лаборатория 1: 0,39; 0,38; 0,46; 0,10; 0,25; 0,25; 0,36.

Лаборатория 2: 0,46; 0,58; 0,52; 0,36; 0,62; 0,68; 0,48; 0,44; 0,32.

Необходимо проверить гипотезу  $H_0: (\mu_X = \mu_Y = \mu)$ , т. е. о том, что математические ожидания совокупностей результатов анализа в лабораториях не отличаются. При этом априорная информация о форме распределений отсутствует. Для решения задачи выборки упорядочивают по возрастанию и располагают в одном ряду:

$$0,10^1; 0,25^1; 0,25^1; 0,32^2; 0,36^2; 0,36^1; 0,38^1; 0,29^1; 0,44^2; 0,46^1; \\ 0,46^2; 0,48^2; 0,52^2; 0,58^2; 0,62^2; 0,68^2. \quad (2.47)$$

Верхние индексы показывают, в какой лаборатории получен результат.

В наших выборках есть две пары не различающихся результатов: 0,36 и 0,36; 0,46 и 0,46. Таких результатов, вообще говоря, может быть сколько угодно много. При определении порядка их расположения в общем ряду руководствуются следующими правилами.

- Если имеется одна пара совпадающих значений, порядок их расположения выбирают случайным образом (бросая монету или с помощью таблицы случайных чисел).
- Если таких пар две, то первую пару располагают случайным образом, а вторую — в обратном порядке (как это сделано нами выше; для значений 0,36 первым поставлен результат из второй выборки, для значений 0,46 — наоборот).
- Если не различающихся (например, по 0,30) результатов в одной выборке имеется три-два результата, в другой — один, то их располагают так

	0,30			или		0,30			0,30
	0,30					0,30			0,30
Если четыре									
	0,30		0,30			0,30			0,30
0,30		и		0,30	или	0,30		и	0,30
0,30				0,30				0,30	0,30
	0,30		0,30			0,30			0,30

и т. д. Общий принцип расположения результатов измерений состоит в том, чтобы не давать приоритета ни одной из выборок: одинаковые результаты измерений из обеих выборок должны быть как можно более равномерно «перемешаны». Естественно, все сказанное не относится к одинаковым результатам из одной выборки — они неразличимы<sup>30</sup>.

При расчете критерия  $U$ , как и при использовании любых непараметрических методов, существенны не сами результаты, а порядок их расположения. Обозначим все результаты, полученные в первой лаборатории, через  $X$ , а во второй – через  $Y$ . Тогда упорядоченный ряд будет выглядеть так:  $XXXYXXYYYY$ . Будем считать «идеальным» такое расположение результатов измерений, при котором сначала расположены все результаты из первой выборки, а затем – из второй. В нашем случае –  $XXXXXXXXYY$ . Дальнейший анализ состоит в подсчете «нарушений» (*инверсий*) расположения результатов по сравнению с их «идеальным» расположением. Одной инверсией считается такое расположение, когда впереди какого-либо результата из первой выборки (первой лаборатории) в упорядоченном ряду (2.45) стоит один

<sup>30</sup> Встречающаяся иногда в литературе рекомендация исключать пары не различающихся результатов из разных выборок некорректна — это приводит к потере информации и снижению эффективности критерия.

результат из второй выборки. Если перед некоторым результатом из первой выборки стоят два результата из второй, это считается за две инверсии, и т. д.

Для наглядности изобразим наш ряд значений в форме табл. 2.4, в каждой строке которой помещен только один результат (полученный либо в первой, либо во второй лаборатории). В правом столбце показано число инверсий  $U$ , соответствующее каждому из значений<sup>31</sup>.

Таблица 2.4. Расчет числа инверсий  $U$

Результаты измерения концентрации серы, %		Число инверсий $U$
Лаборатория 1 (выборка 1)	Лаборатория 2 (выборка 2)	
0,10		0
0,25		0
0,25		0
	0,32	—
	0,36	—
0,36		2
0,38		2
0,39		2
	0,44	—
0,46		3
	0,46	—
	0,48	—
	0,52	—
	0,58	—
	0,62	—
	0,68	—
<b>Всего</b>		<b>9</b>

Для проверки гипотезы  $H_0: (\mu_x = \mu_y = \mu)$  найденное общее число инверсий  $U$  сравнивают с табличным значением максимального числа инверсий  $U(\bar{P}, N_1, N_2)$ , при котором гипотеза отвергается ( $\bar{P}$  — принятый уровень доверительной вероятности,  $N_1$  и  $N_2$  — числа результатов измерений в сравниваемых выборках).

Гипотезу  $H_0$  отвергают при  $U \leq U(\bar{P}, N_1, N_2)$ . Таблица для  $U(\bar{P}, N_1, N_2)$  для доверительной вероятности 0,95 приведена в Приложении 7 (таблицы для других уровней  $\bar{P}$  даны, например, в [24]). Для нашего примера  $U(0,95, 7, 9) = 15$ ,  $U = 9$  и, следовательно, гипотеза о равенстве математических ожиданий результатов анализа в лабораториях отвергается при  $\bar{P} = 0,95$ : лаборатории различаются.

Критерий  $U$ , как и другие непараметрические критерии, нечувствителен к наличию грубых промахов (в нашем примере — значение 0,10 в первой выборке).

Как сказано выше, для использования непараметрических критериев не нужно знать (предполагать) форму распределения — в этом их преимущество. Однако если форма распределения известна, они менее эффективны, чем параметрические. Так, если при проверке гипотезы о равенстве двух математических ожиданий мы знаем, что распределение совокупностей нормальное, то критерий  $U$  менее эффективен, чем  $t$ -критерий Стьюдента (то есть по критерию Стьюдента можно выявить меньшее различие между математическими ожиданиями).

<sup>31</sup> В любом упорядоченном ряду вида (2.42) число инверсий можно подсчитывать как относительно первой, так и второй выборки. Сумма инверсий в обоих случаях одинакова.

### 2.5.4. Гипотеза о равенстве математических ожиданий при связанных выборках

На практике часто реализуется ситуация, когда имеются две выборки, содержащие одинаковое число результатов измерений, относительно которых заведомо известно, что они попарно взаимосвязаны в том смысле, что каждая пара измерений выполнена для одного объекта.

**Пример.** Некоторый набор (разных) проб проанализирован двумя лаборантами, причем каждый лаборант выполнил по одному измерению для каждой пробы.

Хотя каждое измерение при этом представляет собой особую выборку, результаты, полученные каждым лаборантом, также условно рассматривают как выборку, а обе выборки называют *связанными*.

Предположим, мы имеем  $m$  пар взаимосвязанных результатов  $(X_i, Y_i)$ , где  $X_i$  — результаты измерений из первой выборки,  $Y_i$  — результаты измерений из второй выборки для одного и того же  $i$ -го объекта. В нашем примере  $X_i$  и  $Y_i$  — результаты измерений, полученные двумя лаборантами для одной и той же  $i$ -й пробы, и число проб равно  $m$ . Понятно, что в отсутствие систематических погрешностей математическое ожидание результатов для каждой пары результатов одно и то же.

Чаще всего в метрологии и аналитической химии по отношению к связанным выборкам проверяют гипотезу  $H_0$  о том, что различия между парами результатов измерений в совокупности обусловлены только случайными причинами, а систематические эффекты отсутствуют. В нашем примере это эквивалентно гипотезе о том, что оба лаборанта работают одинаково, а различия в результатах анализа обусловлены случайными причинами.

Для решения подобных задач используют как парный  $t$ -критерий Стьюдента, так и непараметрические методы.

Использование парного  $t$ -критерия возможно в предположении о нормальном распределении совокупностей. В нашем примере это означает, что если бы для любой из проб одним лаборантом было бы проведено не одно, а бесконечное множество измерений, то они имели бы одно и то же нормальное распределение. Расчеты с помощью парного  $t$ -критерия проводят следующим образом.

Находят разности  $d_i = X_i - Y_i$ . В сделанных предположениях величины  $d_i$  имеют нормальное распределение с математическим ожиданием  $\mu_d$ . Далее проверяют гипотезу  $H'_0$ : ( $\mu_d = 0$ ). Если основная гипотеза  $H_0$  верна, разности  $d_i$  должны быть беспорядочно рассеяны вокруг нулевого значения, т. е. гипотезы  $H_0$  и  $H'_0$  эквивалентны. Для проверки гипотезы рассчитывают величину

$$t = \frac{|\bar{d}|}{s_d} \sqrt{m}, \quad (2.48)$$

где  $\bar{d}$  и  $s_d$  — соответственно среднее значение и стандартное отклонение величин  $d_i$ .

Величина  $t$  имеет  $t$ -распределение с  $f = m - 2$  степенями свободы. Соответственно проверка гипотез  $H'_0$  и  $H_0$  сводится к сравнению найденного значения  $t$  с процентной точкой  $t$ -распределения  $t(\bar{P}, f)$ , где  $\bar{P}$  — доверительная вероятность,  $f = m - 2$ . Проверяемая гипотеза отвергается при  $t > t(\bar{P}, f)$ .



**Пример.** Для выявления различий в методах определения органической серы в мазуте между двумя лабораториями им был направлен одинаковый набор из  $m = 10$  образцов мазута с разной концентрацией серы. Результаты определения серы в этих образцах (выполненные методом двойного сожжения (в %), приведены в табл. 2.5. В столбце «Лаб. 1 ( $X_i$ )» даны результаты, полученные в первой лаборатории, в столбце «Лаб. 2 ( $Y_i$ )» – во второй. Проверка гипотезы проводится с  $P = 0,95$ .

Таблица 2.5. Результаты определения серы в 10 образцах мазута в разных лабораториях

№ образца ( $i$ )	Результаты анализов (в %)		$d_i = X_i - Y_i$
	Лаб. 1 ( $X_i$ )	Лаб. 2 ( $Y_i$ )	
1	0,55	0,49	+0,06
2	0,14	0,10	+0,04
3	0,44	0,42	+0,02
4	0,66	0,59	+0,07
5	0,93	0,93	0,00
6	0,52	0,48	+0,04
7	0,26	0,23	+0,03
8	0,34	0,35	–0,01
9	0,79	0,76	+0,03
10	0,45	0,41	+0,04

В столбце « $d_i = X_i - Y_i$ » – разности  $X_i$  и  $Y_i$ . Находим  $\bar{d} = +0,32$ ,  $s = 0,0244$  и  $t = \frac{|\bar{d}|}{s} \sqrt{m} = 41,47$ . Табличное значение  $t(0,975, 8) = 2,306$ , то есть гипотеза о равенстве математических ожиданий отвергается с  $\bar{P} = 0,95$ . Заметим, что нулевая гипотеза отвергается и с  $\bar{P} = 0,99$ :  $t(0,995, 8) = 3,355$ .

Применение парного  $t$ -критерия ограничено, так как для проверки предположения о нормальном распределении требуется провести специальный достаточно объемный эксперимент. Если уверенности в нормальном распределении результатов нет, целесообразно использовать непараметрические методы – критерий знаков и парный критерий Вилкоксона.

**Критерий знаков.** Простейшим из непараметрических критериев, используемых для случая связанных выборок, является критерий знаков. Найдем для каждой пары результатов измерений ( $X_i, Y_i$ ) знак разности  $d_i = X_i - Y_i$ :  $\{+\}$  при  $d_i > 0$  и  $\{-\}$  при  $d_i < 0$ . Очевидно, если обе связанные выборки получены из одной совокупности, вероятность появления знака  $\{+\}$  равна  $P = 0,5$ . Таким образом, необходимо проверить гипотезу  $H_0$ : ( $P = 0,5$ ). В качестве альтернативной чаще всего выбирают гипотезу (двустороннюю)  $H_1$ : ( $P \neq 0,5$ ).

Для проверки гипотезы  $H_0$  подсчитывают число менее часто встречающихся знаков  $N_{\min}$  и сравнивают его с критическим значением  $N_{\text{кр}}(\bar{P}, m)$ , где  $\bar{P}$  – доверительная вероятность. Значения приведены в Приложении 8. Гипотеза  $H_0$  отвергается, если  $N_{\min} \leq N_{\text{кр}}(\bar{P}, m)$ .

**Пример.** Применим критерий знаков к данным из табл. 2.8. Среди величин  $d_i$  8 положительных и 1 отрицательная. В Приложении 7 для  $m = 10$  для  $\bar{P} = 0,95$  находим критическое значение, равное 1, то есть нулевая гипотеза отвергается. Заметим, что, в отличие от случая

применения парного  $t$ -критерия, нулевая гипотеза не отвергается с  $\bar{P} = 0,99$ . Это связано с очень низкой мощностью критерия знаков.

Способ проверки по критерию знаков при односторонней альтернативе  $H_1$ : ( $P > 0,5$ ) или  $H_1$ : ( $P < 0,5$ ) приведен в [23].

Критерий знаков очень надежен но, к сожалению, хорошо работает лишь при большом ( $m > 20$ ) числе измерений.

**Парный критерий Вилкоксона.** Если критерий знаков не выявил достоверного различия между выборками, особенно при небольшом числе измерений в сравниваемых выборках ( $m \leq 20$ ), рекомендуется провести дополнительную проверку с использованием **парного критерия Вилкоксона**. Для этого:

- рассчитывают разности  $d_i = X_i - Y_i$  и располагают их по порядку возрастания;
- присваивают каждой величине  $d_i$  **ранговый номер (ранг)** – ее порядковый номер в получившемся ряду. Если две или более разности  $d_i$  равны между собой, каждой из них присваивают средний ранг: например, если две равные разности  $d_i$  оказались на пятой и шестой позициях в возрастающем ряду, каждой из них присваивают ранговый номер 5.5;
- рассчитывают сумму ранговых номеров для величин  $d_i$ , имеющих отрицательное значение;
- располагают величины  $d_i$  в порядке убывания и присваивают им ранговый номер в получившемся ряду;
- рассчитывают сумму ранговых номеров в получившемся ряду для величин  $d_i$  имеющих положительное значение;
- выбирают меньшую из двух получившихся сумм ранговых номеров и обозначают ее как  $T$ ;
- сравнивают найденную величину  $T$  с критическим значением  $T_{кр}(\bar{P}, m)$  ( $\bar{P}$  – доверительная вероятность,  $m$  – число пар результатов  $(X_i, Y_i)$ ). Гипотеза  $H_0$  отвергается при  $T \leq T_{кр}(\bar{P}, m)$ . Величины  $T_{кр}(\bar{P}, m)$  приведены в Приложении 9.

**Пример.** Применим парный критерий Вилкоксона к данным табл. 2.8. Расположив величины  $d_i$  по возрастанию, получаем ряд:

Значение	-0,01	0,00	+0,02	+0,03	+0,03	+0,04	+0,04	+0,04	+0,06	+0,07
Ранг	1	2	3	4,5	4,5	7	7	7	9	10

Присваиваем ранги и видим, что только одно значение отрицательно. Сумма рангов отрицательных величин равна 1.

Затем расположим величины  $d_i$  по убыванию. Получаем ряд:

Значение	+0,07	+0,06	+0,04	+0,04	+0,04	+0,03	+0,03	+0,02	0,00	-0,01
Ранг	1	2	4	4	4	6,5	6,5	8	9	10

Присваиваем ранги и находим сумму рангов положительных величин равной 36.

Меньшая из получившихся сумм рангов равна 1. Сравниваем ее с  $T_{кр}(\bar{P}, 10)$  из Приложения 8 которые для  $m = 10$  равны 10 для  $\bar{P} = 0,95$  и 5 для  $\bar{P} = 0,95$ . В обоих случаях нулевая гипотеза отвергается, что демонстрирует большую эффективность парного критерия Вилкоксона по сравнению с критерием знаков.

Если  $m > 20$ , можно также рассчитать величину

$$\tilde{T} = \frac{T - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}}, \quad (2.49)$$

нормально распределенную с параметрами  $(0, 1)$ . Проверяемая гипотеза отвергается при  $\tilde{T} > 1,96$  (для  $\bar{P} = 0,95$ ) или при  $\tilde{T} > 2,56$  (для  $\bar{P} = 0,99$ ).

## 2.6. Сложение погрешностей и дисперсионный анализ

Рассмотренные в разд. 2.5 гипотезы касаются ситуаций, в которых исследуются различия между двумя выборками. Однако при обработке результатов измерений часто возникает необходимость в проверке гипотезы о наличии какого-либо эффекта (эффектов) и получении соответствующих оценок в случаях, когда имеется несколько выборок. Если выборок много, выполняют *дисперсионный анализ*. При этом во многих случаях удастся одновременно выявлять влияние нескольких факторов.

Вначале рассмотрим простейшие случаи, из которых станет ясна идея дисперсионного анализа, а также используемые в нем способы проверки гипотезы, основанные на законах сложения погрешностей.

### 2.6.1. Сложение погрешностей. Принцип пренебрежений малыми погрешностями

Предположим, что имеется случайная величина  $X$ , и эта величина есть функция нескольких других взаимно независимых случайных величин  $Z_1, Z_2, \dots, Z_N$ :

$$X = \varphi(Z_1, Z_2, \dots, Z_N). \quad (2.50)$$

Если величины  $Z_1, Z_2, \dots, Z_N$  взаимно независимы, связь дисперсий этих случайных величин описывается следующей формулой:

$$\sigma_X^2 = \left( \frac{d\varphi}{dZ_1} \right)^2 \cdot \sigma_{Z_1}^2 + \left( \frac{d\varphi}{dZ_2} \right)^2 \cdot \sigma_{Z_2}^2 + \dots + \left( \frac{d\varphi}{dZ_N} \right)^2 \cdot \sigma_{Z_N}^2. \quad (2.51)$$

Рассмотрим здесь простейшие случаи, когда случайная величина есть а) сумма (разность) двух случайных величин (аддитивный случай) или б) произведение либо частное двух случайных величин (мультипликативный случай). Например, результат взвешивания жидкости есть разность масс полной и пустой колбы (аддитивный случай); результат объемного анализа есть произведение объема титранта на чувствительность методики (мультипликативный случай).

Тогда имеем:

$$\text{если } X = Z_1 \pm Z_2, \text{ то } \sigma_X^2 = \sigma_{Z_1}^2 + \sigma_{Z_2}^2; \quad (2.52)$$



$$\text{если } X = Z_1 \cdot Z_2 \text{ или } X = \frac{Z_1}{Z_2}, \text{ то } \left( \frac{\sigma_X}{X} \right)^2 = \left( \frac{\sigma_{Z_1}}{Z_1} \right)^2 + \left( \frac{\sigma_{Z_2}}{Z_2} \right)^2. \quad (2.53)$$

Другими словами, для сумм и разностей (аддитивный случай) суммируются дисперсии абсолютных погрешностей, а для произведений и частных (мультипликативный случай) — дисперсии относительных погрешностей.

Формулы (2.52) и (2.53) легко интерполируются на случай любого числа величин  $Z$ . Для стандартных отклонений имеем:

$$\text{если } X = Z_1 + Z_2 + \dots + Z_N, \text{ то } \sigma_X = \sqrt{\sigma_{Z_1}^2 + \sigma_{Z_2}^2 + \dots + \sigma_{Z_N}^2}; \quad (2.54)$$

$$\text{если } X = \frac{Z_1 \cdot Z_2 \cdot \dots \cdot Z_k}{Z_{k+1} \cdot Z_{k+2} \cdot \dots \cdot Z_N}, \text{ то } \sigma_X = \sqrt{\left( \frac{\sigma_{Z_1}}{Z_1} \right)^2 + \left( \frac{\sigma_{Z_2}}{Z_2} \right)^2 + \dots + \left( \frac{\sigma_{Z_N}}{Z_N} \right)^2}. \quad (2.55)$$

Поскольку выше речь шла о случайных величинах, т. е. о генеральных совокупностях, а квадрат стандартного отклонения (дисперсия) выборки есть оценка дисперсии соответствующей совокупности, выражения (2.52–2.55) приближенно справедливы и для величин  $s_Y^2$ . А именно, если результаты измерений величины  $X - X_1, X_2, \dots, X_N$  — имеют стандартное отклонение  $s_X$  и определяются аддитивно несколькими факторами  $Z_1, Z_2, \dots, Z_N$ , для каждого из которых оценены стандартные отклонения  $s_{Z_1}^2, s_{Z_2}^2, \dots, s_{Z_N}^2$ , то

$$s_X \approx \sqrt{s_{Z_1}^2 + s_{Z_2}^2 + \dots + s_{Z_N}^2}. \quad (2.56)$$

**Пример.** Определяется концентрация некоторого вещества фотометрическим методом, и известно, что погрешность этого измерения определяется погрешностями стадий анализа — взятия навески, растворения вещества и фотометрирования, причем эти погрешности аддитивны. Известно, что погрешность, вносимая на стадии растворения и другими факторами, пренебрежимо мала. Мы можем оценить:

- стандартное отклонение результатов анализа в целом  $s_{\text{общ}}$ , выполнив несколько измерений одной и той же пробы;
- стандартное отклонение взвешивания  $s_{\text{взв}}^2$ , проведя несколько взвешиваний;
- стандартное отклонение фотометрирования  $s_{\text{ф}}^2$ , многократно измерив оптическую плотность раствора какой-либо пробы.

Если наши предположения верны, будет приближенно выполняться равенство

$$s_{\text{общ}} \approx \sqrt{s_{\text{взв}}^2 + s_{\text{ф}}^2}. \quad (2.57)$$

Аналогична ситуация для мультипликативного случая, когда суммируются квадраты относительных стандартных отклонений: если  $X = \frac{Z_1 \cdot Z_2 \cdot \dots \cdot Z_k}{Z_{k+1} \cdot Z_{k+2} \cdot \dots \cdot Z_N}$ , то

$$s_X \approx \sqrt{\left( \frac{s_{Z_1}}{Z_1} \right)^2 + \left( \frac{s_{Z_2}}{Z_2} \right)^2 + \dots + \left( \frac{s_{Z_N}}{Z_N} \right)^2}. \quad (2.58)$$

Для зависимостей, включающих как суммирование, так и умножение (деление) величин, проще всего разбить их на отдельные выражения, для каждого из которых справедливы либо формулы (2.52, 2.54), либо формулы (2.53, 2.55). Например, если

$$X = \frac{(Z_1 + Z_2)}{(Z_3 + Z_4)}, \quad (2.59)$$

то это выражение надо разбить на две части  $(Z_1 + Z_2)$  и  $(Z_3 + Z_4)$ , для каждой из них применить формулу (2.52), а к результатам применить формулу (2.53). В более сложных случаях приходится непосредственно использовать формулу (2.51) — к счастью, это бывает редко.

Формулы сложения погрешностей широко применяются, в частности, при расчете суммарной неопределенности результатов измерений (см. разд. 3.5.2).

Следствием соотношений (2.52, 2.53 и сл.), а также свойств  $\chi^2$ -распределения является принцип пренебрежения малыми погрешностями: если общая погрешность, аддитивно определяется двумя независимыми источниками погрешности, характеризуемыми стандартными отклонениями  $s_1$  и  $s_2$ , и  $s_1/s_2 \leq 0,3$ , то меньшей погрешностью  $s_1$  можно пренебречь, исключив соответствующий источник погрешности из рассмотрения. Поскольку суммируются квадраты стандартных отклонений, вклад меньшей погрешности в общую в этом случае можно не брать в расчет.

Заметим, что систематическими погрешностями пренебрегают, когда они составляют не более 0,1–0,2 общей погрешности [3, 32]; это связано с тем, что суммируются сами систематические погрешности, а не их квадраты. Довольно часто систематические погрешности «трансформируют» в случайные, на которые полностью распространяется принцип пренебрежения малыми погрешностями (см. разд. 2.5.3.2).

Из формул (2.54, 2.56) следует еще одно важное соотношение. Пусть мы имеем выборку  $X_1, X_2, \dots, X_N$  из нормально распределенной совокупности со стандартным отклонением  $s_X$  и число измерений  $N$  кратно целому числу  $k$ . Сгруппируем результаты в группы по  $k$  измерений в каждой и усредним их в пределах каждой группы, обозначив средние значения как  $\bar{Y}_j$ :

Исходная выборка	$\underbrace{X_1, X_2, \dots, X_k}_{\bar{Y}_1}$	$\underbrace{X_{k+1}, X_{k+2}, \dots, X_{2k}}_{\bar{Y}_2}$	$\dots$	$\underbrace{X_{N-k+1}, X_{N-k+2}, \dots, X_N}_{\bar{Y}_{N/k}}$
Средние значения	$\bar{Y}_1$	$\bar{Y}_2$	$\dots$	$\bar{Y}_{N/k}$

Тогда стандартное отклонение новой выборки, состоящей из средних значений  $s_{\bar{Y}}$ , будет равно (приблизительно)

$$s_{\bar{Y}} \approx \frac{s_X}{\sqrt{k}}. \quad (2.60)$$

Для генеральных совокупностей, соответствующих этим выборкам, справедлива точная формула

$$\sigma_{\bar{y}} \approx \frac{\sigma_x}{\sqrt{k}}, \quad (2.61)$$

где  $\sigma_{\bar{y}}$  и  $\sigma_x$  — соответственно, средние квадратичные отклонения совокупности для выборки из средних значений и исходной совокупностей. Соотношения (2.60, 2.61) лежат в основе *дисперсионного анализа*.

### 2.6.2. Простейший (однофакторный) дисперсионный анализ

Используя формулы (2.54, 2.56), можно оценить общую погрешность, зная ее составляющие. Обратная задача решается в рамках *дисперсионного анализа*.

Пример. Для аттестации стандартного образца смазочного масла на содержание серы экземпляры этого образца были разосланы в  $m$  лабораторий ( $m > 3$ ), каждая из которых провела  $N$  измерений содержания серы в этих экземплярах в одних и тех же условиях ( $N > 2$ ). Предположение о равном числе измерений в лабораториях введено для простоты. Таким образом был проведен эксперимент, структура которого показана на рис. 2.10.

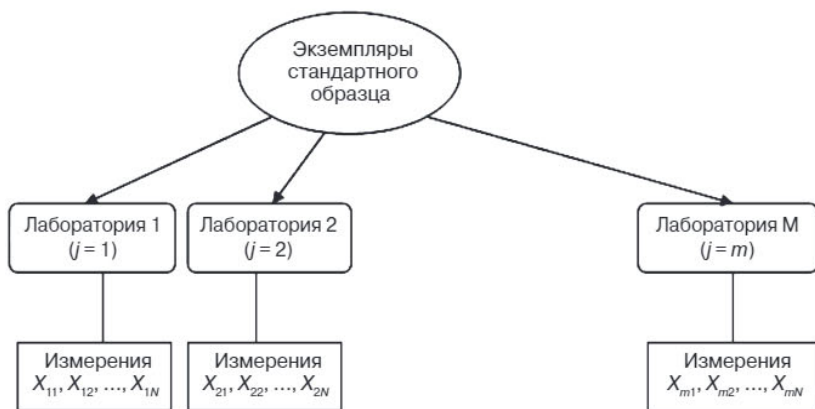


Рис. 2.10. Структура эксперимента по аттестации стандартного образца смазочного масла на содержание серы

Результаты, полученные в каждой  $j$ -й лаборатории, можно рассматривать как отдельную выборку (в дисперсионном анализе называемую группой). Предполагается, что каждой выборке соответствует своя генеральная совокупность. Полученные результаты измерений можно представить в виде табл. 2.6, в которую наряду с результатами измерений  $X_{ji}$  внесены характеристики каждой из выборок:

- среднее значение  $\bar{X}_j = \sum_{i=1}^N X_{ji}$ ;
- стандартное отклонение  $s_j = \sqrt{\frac{\sum_{i=1}^N (X_{ji} - \bar{X}_j)^2}{N-1}}$ ;
- число степеней свободы  $f$  (поскольку число измерений во всех группах одинаково,  $f = N - 1$  для любой группы).



Таблица 2.6. Данные для простейшего (однофакторного) дисперсионного анализа

Номер выборки (группы)	Номер измерения в выборке (группе)						Характеристики выборки			Параметры совокупности	
	1	2	...	$i$	...	$N$	Среднее значение	Станд. отклонение	Число ст. свободы	Математическое ожидание	Стандартное отклонение
1	$X_{11}$	$X_{12}$	...	$X_{1i}$	...	$X_{1N}$	$\bar{X}_1$	$s_1$	$N-1$	$\mu_1$	$s$
2	$X_{21}$	$X_{22}$	...	$X_{2i}$	...	$X_{2N}$	$\bar{X}_2$	$s_2$	$N-1$	$\mu_2$	$s$
...	...	...	...	...	...	...	...	...	...	...	...
$j$	$X_{j1}$	$X_{j2}$	...	$X_{ji}$	...	$X_{jN}$	$\bar{X}_j$	$s_j$	$N-1$	$\mu_j$	$s$
...	...	...	...	...	...	...	...	...	...	...	...
$m$	$X_{m1}$	$X_{m2}$	...	$X_{mi}$	...	$X_{mN}$	$\bar{X}_m$	$s_m$	$N-1$	$\mu_m$	$s$

Кроме того, для совокупностей, соответствующих каждой группе, в таблице указаны математические ожидания  $\mu_j$  и стандартные отклонения  $\sigma$ , предполагаемые одинаковыми для всех совокупностей.

Стандартные предположения, которые делаются при дисперсионном анализе, следующие:

- все выборки получены из нормальных совокупностей с одинаковыми стандартными отклонениями  $\sigma$  и математическими ожиданиями  $\mu_1, \mu_2, \dots, \mu_m$  (см. табл. 2.6);
- результаты, входящие в разные выборки, независимы, а входящие в одну выборку получены в одинаковых условиях (то есть являются случайной выборкой из одной совокупности)<sup>32</sup>.

В рамках сделанных предположений данные могут быть описаны моделью

$$X_{ji} = \alpha + \alpha_j + \varepsilon_{ji}, \quad (2.62)$$

где  $X_{ji}$  — результат  $i$ -го измерения в  $j$ -й группе (выборке),  $\alpha$  — постоянный эффект (математическое ожидание совокупности, соответствующей всем значениям  $X_{ji}$ ),  $\alpha_j$  — эффект выборки (в данном примере это — эффект лаборатории);  $\varepsilon_{ji}$  — случайная погрешность. Эффекты  $\alpha_j$  и  $\varepsilon_{ji}$  имеют нормальное распределение с нулевыми средними значениями и дисперсиями  $\sigma_{\alpha}^2$  и  $\sigma_{\varepsilon}^2$  соответственно.

Нас интересует, различаются ли результаты, полученные в разных лабораториях, т. е. различаются ли они в совокупности. Это означает, что надо проверить гипотезу  $H_0: (\mu_1 = \mu_2 = \dots = \mu_m = \mu)$ , то есть гипотезу о том, что все рассматриваемые выборки получены из одной и той же нормальной совокупности  $X \in N(\mu, \sigma)$ . Это простейшая задача дисперсионного анализа — *однофакторный дисперсионный анализ*.

В терминах модели (2.62) *основную гипотезу простейшего (однофакторного) дисперсионного анализа* можно записать как  $H_0: (\alpha_j = 0)$ , то есть отсутствие эффекта выборки (группы, в нашем примере — лаборатории).

Как сказано выше, важнейшим предположением в дисперсионном анализе является равенство стандартных отклонений совокупностей, из которых взяты

<sup>32</sup> Заметим, что в отличие от первого предположения выполнение второго зависит от организации эксперимента.

рассматриваемые выборки. Эту гипотезу можно проверить с помощью *критерия Бартлета* (разд. 2.5.2). Кроме того, можно применить *критерий Кохрена* (разд. 2.5.2) и, если подозрительно выделяющаяся выборка окажется не принадлежащей к генеральной совокупности, удалить ее из расчетов. Это справедливо для всех вариантов дисперсионного анализа.

В случае, когда гипотеза о равенстве дисперсий отвергается, можно попытаться разделить выборки на классы, в пределах которых дисперсии однородны, и проводить дисперсионный анализ отдельно для каждого класса [33]. Однако этот прием эффективен только при достаточно большом объеме статистического материала.

Для иллюстрации основной идеи дисперсионного анализа примем, как и прежде, что объемы всех выборок одинаковы — по  $N$  измерений в группе (см. табл. 2.3). Тогда, предполагая, что гипотеза  $H_0: (\mu_1 = \mu_2 = \dots = \mu_m = \mu)$  верна, можно найти стандартное отклонение величин  $\bar{X}_j$  (их обычно называют групповыми средними) двумя способами:

- непосредственно по формуле

$$s_{(1)} = \sqrt{\frac{\sum_{j=1}^m (\bar{X}_j - \bar{\bar{X}})^2}{m-1}}, \quad (2.63)$$

где  $\bar{\bar{X}} = \frac{\sum_{j=1}^m \bar{X}_j}{m}$  — общее среднее всех  $X_{ji}$ ;

- с использованием соотношений (2.50, 2.51):

$$s_{(2)} = \frac{s_{\text{общ}}}{\sqrt{N}}, \quad (2.64)$$

где  $s_{\text{общ}} = \sqrt{\frac{\sum_{j=1}^m \sum_{i=1}^N (X_{ji} - \bar{\bar{X}})^2}{mN-1}}$  — общее стандартное отклонение единичных измерений.

Для проверки нашей гипотезы достаточно сравнить стандартные отклонения  $s_{(1)}$  и  $s_{(2)}$  с использованием  $F$ -критерия (см. выше, п. ...). Если

$$F \frac{s_{(1)}^2}{s_{(2)}^2} \quad (2.65)$$

больше критического значения  $F(\bar{P}, f_1 = m-1, f_2 = mN-1)$ , то гипотеза  $H_0: (\mu_1 = \mu_2 = \dots = \mu_m = \mu)$  отвергается с принятым уровнем доверительной вероятности  $\bar{P}$ , т. е. имеется эффект выборки, в нашем примере эффект лаборатории. Другими словами, лаборатории отличаются друг от друга в совокупности.

Приведенный способ дан нами только для наглядной иллюстрации метода дисперсионного анализа и в описанном виде неэффективен. На практике при дисперсионном анализе проверку гипотезы с помощью  $F$ -критерия проводят на уровне единичных измерений, пересчитывая дисперсию групповых средних. При этом стандартное отклонение групповых средних сравнивают не с общим стандартным отклонением, а с усредненным средним стандартным отклонением для всех групп. Покажем, как это делается в варианте однофакторного дисперсионного анализа (по-прежнему предполагая, что выборки одинаковы по объему, то есть для данных из табл. 2.6).

Прежде всего еще раз определим компоненты, из которых складывается общая дисперсия, используя на сей раз стандартные обозначения. Дисперсию, обусловленную рассматриваемым эффектом (эффект выборки или группы, в нашем примере – различиями между лабораториями), обозначим как  $s_{\text{эфф}}^2$ . Это – несмещенная оценка дисперсии соответствующей совокупности  $\sigma_{\alpha}^2$  (см. модель (2.62) и сл.). Дисперсию, определяемую случайным разбросом результатов внутри групп (в нашем примере разброс результатов внутри лабораторий), обозначим как  $s_{\text{ост}}^2$ , она является оценкой дисперсии соответствующей совокупности  $\sigma_{\alpha}^2$  (см. модель (2.62) и сл.). Эту дисперсию называют остаточной, поскольку она «остается» в случае учета остальных погрешностей. И, наконец, кажущуюся дисперсию, соответствующую разбросу результатов между средними выборок (групп), обозначим как  $s_{\text{ср}}^2$ . Этот разброс является результатом взаимодействия (наложения) разброса внутри и между лабораториями. Тогда их дисперсии связаны соотношением

$$s_{\text{ср}}^2 = s_{\text{ост}}^2 + N s_{\text{эфф}}^2. \quad (2.66)$$

Соотношение (2.66) часто называют формулой разложения дисперсий.

Расчетные формулы приведены в табл. 2.7;  $SS$  (Sum of Squares) – стандартное обозначение для суммы квадратов.

**Таблица 2.7.** Расчетные формулы простейшего (однофакторного) дисперсионного анализа

Источник погрешности	Число степеней свободы	Сумма квадратов	Дисперсия	Компоненты дисперсии
Разброс между группами (выборками)	$f_{\text{эфф}} = m - 1$	$s_{\text{эфф}}^2 = \sum_{j=1}^m N (\bar{X}_j - \bar{\bar{X}})^2$	$s_{\text{эфф}}^2 = S_{\text{эфф}} / f_{\text{эфф}}$	$s_{\text{ср}}^2 = s_{\text{ост}}^2 + N s_{\text{эфф}}^2$
Разброс внутри групп (остаточный)	$f_{\text{ост}} = m(N - 1)$	$s_{\text{ост}}^2 = \sum_{j=1}^m \sum_{i=1}^N (X_{ji} - \bar{X}_j)^2$	$s_{\text{ост}}^2 = S_{\text{ост}} / f_{\text{ост}}$	
Общий разброс	$f = mN - 1$	$S = S_{\text{эфф}} + S_{\text{ост}} = \sum_{j=1}^m \sum_{i=1}^N (X_{ji} - \bar{\bar{X}})^2$		

Форма таблицы стандартна и соответствует, в частности, форме выдачи результатов большинством компьютерных программ.

Для проверки гипотезы  $H_0: (\mu_1 = \mu_2 = \dots = \mu_m = \mu)$ , которая может быть записана и в форме  $H_0: (\alpha_j = 0)$  рассчитывают отношение

$$F = \frac{s_{\text{эфф}}^2}{s_{\text{ост}}^2}. \quad (2.67)$$

Гипотеза  $H_0$  отвергается при  $F > F(\bar{P}, f_{\text{эфф}}, f_{\text{ост}})$ , где  $F(\bar{P}, f_{\text{эфф}}, f_{\text{ост}})$  – процентная точка  $F$ -распределения при принятой доверительной вероятности  $\bar{P}$  и числах степеней свободы  $f_{\text{эфф}}$  и  $f_{\text{ост}}$ .



### 2.6.3. Другие варианты дисперсионного анализа

Основной принцип дисперсионного анализа един для всех его вариантов: рассчитываются суммы квадратов и оценки дисперсии, соответствующие рассматриваемым эффектам, и гипотеза об отсутствии того или иного эффекта проверяется по  $F$ -критерию. Выделим несколько частных случаев.

Двухфакторный дисперсионный анализ с одним наблюдением в ячейке возможен в том случае, когда результаты эксперимента могут быть записаны в виде табл. 2.8.

Таблица 2.8. Данные для двухфакторного дисперсионного анализа с одним наблюдением в ячейке

Уровни фактора А	Уровни фактора В						Среднее групповое (по строке)
	1	2	...	$i$	...	$N$	
1	$X_{11}$	$X_{12}$	...	$X_{1i}$	...	$X_{1N}$	$\bar{X}_{1\bullet}$
2	$X_{21}$	$X_{22}$	...	$X_{2i}$	...	$X_{2N}$	$\bar{X}_{2\bullet}$
...	...	...	...	...	...	...	...
$j$	$X_{j1}$	$X_{j2}$	...	$X_{ji}$	...	$X_{jN}$	$\bar{X}_{j\bullet}$
...	...	...	...	...	...	...	...
$M$	$X_{M1}$	$X_{M2}$	...	$X_{Mi}$	...	$X_{MN}$	$\bar{X}_{M\bullet}$
Среднее групповое (по столбцу)	$\bar{X}_{\bullet 1}$	$\bar{X}_{\bullet 2}$	...	$\bar{X}_{\bullet i}$	...	$\bar{X}_{\bullet N}$	

Такую структуру данных имеют, например, результаты исследования зависимости интенсивности линии испускания хрома от мощности горелки (фактор А) и скорости распыления раствора (фактор В) при определении хрома методом эмиссионной спектроскопии с индуктивно связанной плазмой. В этом случае уровни фактора А – разные уровни мощности (им соответствует индекс  $j$ ), а уровни фактора В – разные скорости распыления (им соответствует индекс  $i$ ), причем уровни факторов могут быть пронумерованы в любом порядке. При каждом сочетании уровней  $i$  и  $j$  проводится одно измерение интенсивности и его результат  $X_{ji}$  (одно наблюдение в ячейке). Значок «•» здесь означает усреднение по индексу, на месте которого он стоит, например  $\bar{X}_{\bullet N} = \sum_{j=1}^M X_{jN} / M$ .

Эти данные могут быть описаны моделью двухфакторного дисперсионного анализа

$$X_{ji} = \mu + \alpha_j + \beta_i + \varepsilon_{ji}, \quad (2.68)$$

где  $X_{ji}$  – результат измерения при  $j$ -м уровне фактора А и  $i$ -м уровне фактора В;  $\mu$  – генеральное среднее всех мыслимых результатов измерений;  $\alpha_j$  и  $\beta_i$  – изучаемые эффекты;  $\varepsilon_{ji}$  – случайная погрешность. Обычно предполагают нормальное распределение случайной погрешности и независимость результатов измерений при различных сочетаниях уровней А и В. Проверяют гипотезы  $H_A$ : ( $\alpha_j = 0$ ) и  $H_B$ : ( $\beta_i = 0$ ), то есть гипотезы об отсутствии влияния факторов А и В на результаты измерений.

Для проверки этих гипотез рассчитывают:

– факторную дисперсию для фактора А:

$$s_A^2 = \frac{\sum_{j=1}^M N(\bar{X}_{j\bullet} - \bar{\bar{X}})^2}{M-1}, \quad (2.69)$$

где  $\bar{\bar{X}}$  – общее среднее всех значений  $\bar{\bar{X}} = \sum_{j=1}^M \sum_{i=1}^N X_{ji} / MN$ ,  $\bar{X}_{j\bullet}$  – групповое среднее по  $j$ -й строке;

– факторную дисперсию для фактора В:

$$s_B^2 = \frac{\sum_{i=1}^N M(\bar{X}_{\bullet i} - \bar{\bar{X}})^2}{N-1}, \quad (2.70)$$

где  $\bar{X}_{\bullet i}$  – групповое среднее для  $i$ -го столбца.

Остаточную дисперсию рассчитывают по формуле

$$s_0^2 = \frac{\sum_{j=1}^M \sum_{i=1}^N (X_{ji} - \bar{X}_{j\bullet} - \bar{X}_{\bullet i} + \bar{\bar{X}})^2}{(M-1)(N-1)}. \quad (2.71)$$

Для проверки гипотезы  $H_A: (\alpha_j = 0)$  рассчитывают величину  $F_A = s_A^2/s_0^2$ . Если  $F_A > F(\bar{P}, f_1, f_2)$ , где  $F(\bar{P}, f_1, f_2)$  – процентная точка  $F$ -распределения при принятой доверительной вероятности  $\bar{P}$  и числах степеней свободы  $f_1 = M-1, f_2 = (M-1)(N-1)$ , то гипотезу  $H_A$  отвергают.

Аналогично проверяют гипотезу  $H_B: (\beta_i = 0)$ . Находят  $F_B = s_B^2/s_0^2$  и гипотезу  $H_B$  отвергают при  $F_B > F(\bar{P}, f_1, f_2)$ , где  $\bar{P}$  – принятая доверительная вероятность,  $f_1 = N-1, f_2 = (M-1)(N-1)$ .

Как видно, принцип дисперсионного анализа неизменен – по критерию сравниваются факторная и остаточная дисперсии.

Аналогично выполняется *многофакторный дисперсионный анализ с одним испытанием в ячейке*.

Двухфакторный дисперсионный анализ с несколькими испытаниями в ячейке применяется, когда для каждого сочетания уровней факторов А и В («в одной ячейке») проводится более одного измерения. В этом случае способ проверки гипотез  $H_A$  и  $H_B$  отличается от описанного выше лишь тем, что при расчете дисперсий используются средние значения результатов, находящиеся в каждой ячейке, а остаточная дисперсия оценивается как среднее стандартных отклонений во всех ячейках. Это верно и для *многофакторного дисперсионного анализа с несколькими испытаниями в ячейке*.

При обработке результатов измерений часто используется *гнездовой план эксперимента (nested design)*, когда каждому уровню одного из факторов соответствуют несколько уровней другого фактора, каждому из которых в свою очередь соответствуют уровни третьего фактора и т. д. Такова, например, структура межлабораторного эксперимента (см. разд. 3.6.3.1), когда один образец сравнения анализируется несколькими лабораториями, каждая из которых проводит несколько серий экспериментов и в каждой серии выполняется несколько параллельных измерений. В таких случаях дисперсионный анализ может быть сведен к последовательному

применению техники однофакторного дисперсионного анализа. Эти и более сложные случаи дисперсионного анализа (в частности, в модели может учитываться взаимодействие факторов) рассмотрены в монографии [33].

Отметим, что иногда при расчете факторной дисперсии получаются отрицательные значения, что соответствует мнимому стандартному отклонению. Такая ситуация не имеет физического смысла и возникает при невыполнении основных предположений дисперсионного анализа.

#### 2.6.4. Несколько замечаний о планировании эксперимента для последующего дисперсионного анализа

В большинстве случаев план эксперимента можно выбирать. Теории оптимального построения плана эксперимента хорошо разработаны [34]; мы не будем рассматривать их подробно. Ограничимся несколькими замечаниями.

В тех случаях, когда остаточной дисперсией (случайной погрешностью измерений) можно пренебречь, используют полные или дробные факторные планы. В приложении к анализу вещества они рассмотрены в [20]. Строго говоря, в этом случае дисперсионный анализ не нужен.

Если остаточной дисперсией пренебречь нельзя и для обработки результатов необходимо применять дисперсионный анализ, *эксперименты фактически не поддаются строгому планированию до тех пор, пока неизвестны погрешности*, т. е. до проведения хотя бы прикидочных опытов. Именно это привело к резкому падению интереса к планированию эксперимента при химическом анализе в последние годы.

Тем не менее можно сформулировать несколько следующих простых рекомендаций.

Необходимо стремиться к «сбалансированному» плану эксперимента, т. е. к равному числу измерений в ячейке на каждом уровне группировки. Например, для рассмотренного в разд. 3.6.3.1 межлабораторного эксперимента с гнездовым планом эффективнее, чтобы все лаборатории проводили одинаковое число серий, выполняя в каждой серии одинаковое число измерений. При сильно несбалансированных планах эксперимента приходится применять специальные методы расчета [35].

Если план эксперимента не диктуется экономическими или другими соображениями, лучше при сохранении общего числа измерений иметь больше уровней для основного оцениваемого эффекта. Так, для рассмотренного выше в разд. 2.6.1 примера аттестации стандартного образца (однофакторный дисперсионный анализ) лучше иметь 50 лабораторий, выполняющих по два измерения, чем 10 лабораторий, выполняющих по десять измерений.

## 2.7. Корреляционный и регрессионный анализ

Одна из задач, постоянно встречающихся при обработке результатов измерений, — поиск зависимости между двумя величинами. В частности, эта задача возникает при построении градуировочной зависимости при косвенных измерениях (см. разд. 1.3.1). Как уже отмечалось, именно такие измерения и проводятся при количественном химическом анализе вещества.



Чаще всего в практике измерений приходится иметь дело с задачей нахождения зависимости между двумя величинами, причем эта зависимость исходно полагается линейной. Основные методы решения такой задачи мы и рассмотрим; некоторые замечания по поводу более сложных случаев см. в разд. 2.7.2.3.

### 2.7.1. Проверка взаимосвязи двух величин

Предположим, имеется  $m$  пар значений  $(X_i, Y_i)$ , причем  $X_i$  — результаты измерений величины  $X$ , а  $Y_i$  — результаты измерений величины  $Y$ <sup>33</sup>. Первый вопрос, на который обычно ищут ответ, — имеется ли вообще связь между этими величинами. Другими словами, необходимо проверить гипотезу  $H_0$ : величины  $X$  и  $Y$  независимы.

Проверка взаимосвязи двух величин не представляет никакой трудности в случае, когда случайные погрешности малы, а диапазон, в котором лежат результаты измерений, достаточно широк. Тогда наличие или отсутствие связи и ее вид становятся очевидными, если нанести точки  $X_i, Y_i$  на координатную плоскость (рис. 2.11). Во всех остальных случаях применяют расчетные методы проверки основной гипотезы.

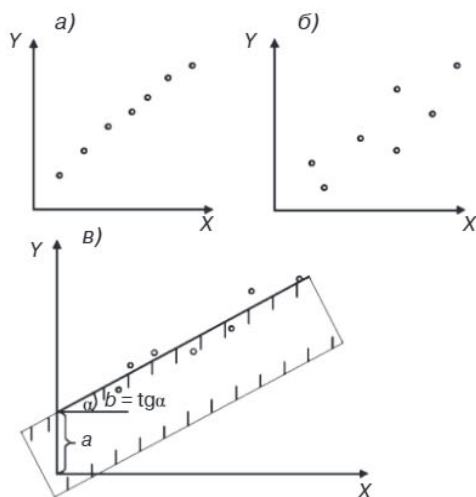


Рис. 2.11. Графическое представление зависимости двух величин: очевидная линейная зависимость (а); явная зависимость отсутствует (б); использование прозрачной линейки для нахождения уравнения линейной зависимости (в)

Весьма универсальным является *графический метод с использованием непараметрического критерия* [20]. По этому методу выполняют следующие операции (рис. 2.12):

- наносят точки  $(X_i, Y_i)$  на координатную плоскость (всего  $m$  точек, в нашем примере  $m = 11$ );

<sup>33</sup> Статистические методы, рассматриваемые ниже, не зависят от способа получения значений  $X_i, Y_i$ . Это могут быть результаты измерений двух величин, результаты измерений одной и той же величины разными методами, расчетные значения и т. д.

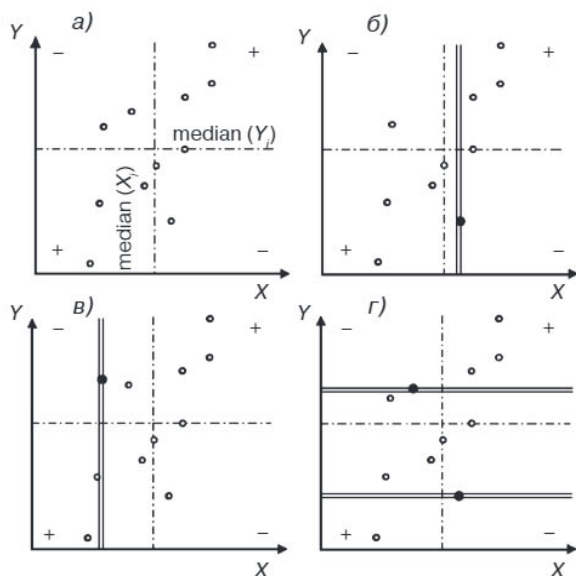


Рис. 2.12. Непараметрический метод выявления взаимосвязи двух величин

- наносят точку с координатами, соответствующими медианам величин  $X_p$  и  $Y_p$ , то есть с координатами  $(median(X_p), median(Y_p))$ , и переносят в эту точку начало координат;
- приписывают знак «+» первому и третьему квадрантам и знак «-» второму и четвертому квадрантам получившейся системы координат (рис. 2.12а);
- перемещают линейку, расположенную параллельно оси  $Y$ , вправо от центра новых координат до тех пор, пока в первом или четвертом квадранте не останется одна точка ( $\bullet$ ). Наложив край линейки на эту точку, подсчитывают число точек  $n_p$ , оставшихся справа от линейки в другом открытом квадранте. Например, если эта точка оказалась в четвертом квадранте (как показано на рис. 2.12б), подсчитывают число точек, оставшихся справа от линейки в первом квадранте ( $n_p = 4$ );
- повторяют то же, перемещая линейку: а) параллельно оси  $Y$  влево от центра новых координат (рис. 2.12в),  $n_p = 2$ ; б) параллельно оси  $X$  вверх от центра новых координат (рис. 2.9г),  $n_p = 3$ ; в) параллельно оси  $X$  вниз от центра новых координат (рис. 2.12г),  $n_p = 1$ ;
- рассчитывают алгебраическую сумму  $N_p$  полученных величин  $n_p$  с учетом знака квадранта, в котором лежали подсчитанные точки:

$$N_p = \sum_1^4 n_p. \quad (2.72)$$

В нашем примере эта сумма равна  $N_p = 4 + 2 + 3 + 1 = 10$  (поскольку все точки оказались в квадрантах, помеченных знаком «+»).

- сравнивают абсолютное значение полученной величины  $N_p$  с критическим значением  $N_p(\bar{P})$  из табл. 2.5.



Таблица 2.9. Критические значения  $N_p(\bar{P})$  для непараметрического критерия

$\bar{P}$	0,90	0,95	0,98	0,99
$N_p(\bar{P})$	9	11	13	14

Нулевая гипотеза отвергается при  $|N_p| \leq N_p(\bar{P})$ . В нашем примере гипотеза об отсутствии взаимосвязи величин  $X$  и  $Y$  отвергается при принятой доверительной вероятности  $\bar{P} = 0,95$ , так как  $N_p = 10$  меньше  $N_p(\bar{P} = 0,95) = 11$ .

Как и другие непараметрические методы, рассмотренный тест не предполагает известного распределения величин  $X$  и  $Y$ , однако хорошо работает либо при больших  $m$ , либо при линейной зависимости и небольших погрешностях измерений.

В случае нормально распределенных величин  $X$  и  $Y$  проверку гипотезы об отсутствии линейной зависимости вида  $Y = aX + b$ , а также оценку силы связи между ними проводят с использованием *выборочного коэффициента линейной корреляции*  $r$  (часто называемого просто коэффициентом корреляции).

Взаимозависимость двух/величин  $X$  и  $Y$  по результатам  $m$  их измерений ( $X_i, Y_i$ ) может быть охарактеризована *выборочной ковариацией*  $s_{XY}$

$$s_{XY} = \frac{\sum_{i=1}^m (X_i - \bar{X})(Y_i - \bar{Y})}{m-1}. \quad (2.73)$$

Выборочная ковариация тем больше, чем «синхроннее» изменяются результаты измерений обеих величин. Если эти результаты находятся «в противофазе» (чем больше значение  $X_i$ , тем меньше значение  $Y_i$ ), то ковариация будет отрицательной. Именно поэтому выборочная ковариация хорошо отражает взаимосвязь величин  $X$  и  $Y$ .

Однако непосредственное использование выборочной ковариации неудобно, поскольку ее значение зависит от величин и размерностей  $X$  и  $Y$ . На практике применяют *выборочный коэффициент линейной корреляции*  $r_{XY}$ :

$$r_{XY} = \frac{s_{XY}}{s_X s_Y}, \quad (2.74)$$

где  $s_X$  и  $s_Y$  — стандартные отклонения величин  $X$  и  $Y$ :

$$s_X = +\sqrt{\sum_{i=1}^m \frac{(X_i - \bar{X})^2}{m-1}} \text{ и } s_Y = +\sqrt{\sum_{i=1}^m \frac{(Y_i - \bar{Y})^2}{m-1}}.$$

Полученная безразмерная величина  $r_{XY}$  (в публикациях ее часто называют просто «коэффициент корреляции», что не совсем корректно, и обозначают как  $r$ , если из текста понятно, связь каких именно величин оценивается) лежит между  $-1$  и  $+1$  ( $-1 \leq r_{XY} \leq +1$ ). Она близка к нулю при отсутствии связи между измеряемыми величинами  $X$  и  $Y$  и увеличивается, приближаясь к  $+1$  при наличии положительной связи между ними (соответственно, приближается к  $-1$  при наличии отрицательной связи).

Выборочная ковариация и коэффициент линейной корреляции являются несмещенными оценками одноименных параметров соответствующих совокупностей.

Заметим, что сама по себе величина выборочного линейного коэффициента корреляции не так информативна, как часто считают. Она в значительной степени зависит от соотношения случайных погрешностей измерений и выбранного диапазона измеряемых величин. Если диапазон, в котором лежат измеряемые величины, невелик, то коэффициент корреляции может оказаться близким к нулю даже при наличии линейной зависимости между измеряемыми величинами.

Для проверки гипотезы об отсутствии линейной зависимости между измеряемыми величинами (ее здесь можно записать как  $H_0: (\rho_{XY}=0)$ , где  $\rho_{XY}$  — коэффициент корреляции между случайными величинами  $X$  и  $Y$ , см. разд. 2.3.1.2) рассчитывают величину

$$t = |r_{XY}| \sqrt{\frac{m-2}{1-r_{XY}^2}}. \quad (2.75)$$

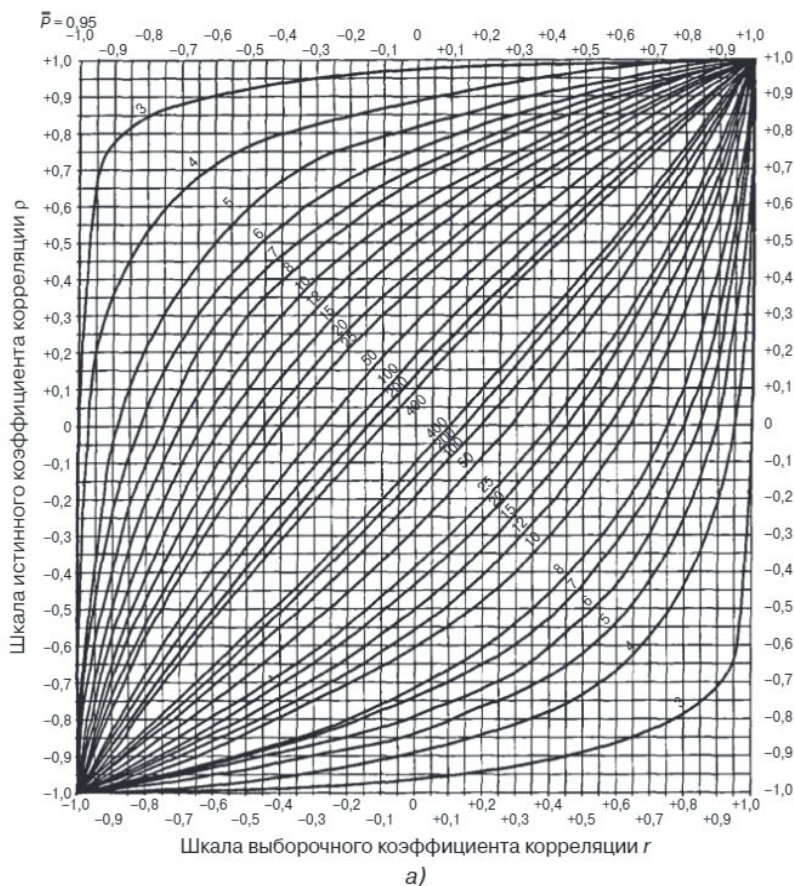


Рис. 2.13. Номограммы для нахождения доверительных интервалов для коэффициентов линейной корреляции ( $\bar{P} = 0,95$ ) (а)



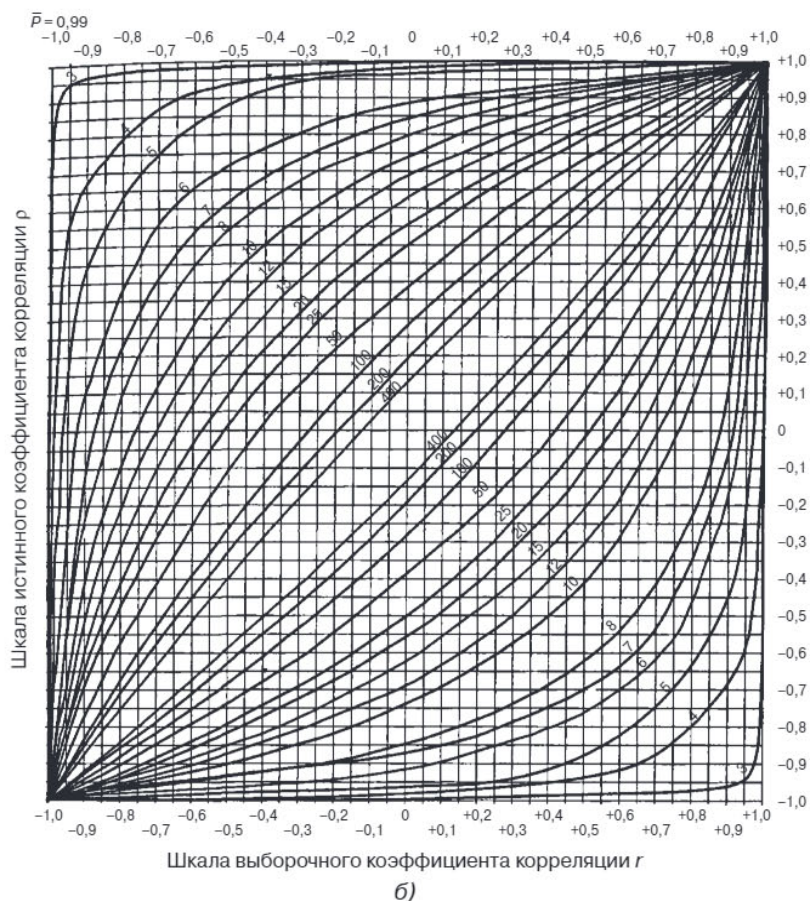


Рис. 2.13 (окончание). Номограммы для нахождения доверительных интервалов для коэффициентов линейной корреляции ( $\bar{P} = 0,99$ ) (б)

Эта величина отвечает  $t$ -распределению Стьюдента с  $f = m - 1$  степенями свободы. Соответственно, гипотеза  $H_0: (\rho_{XY} = 0)$  отвергается с выбранным уровнем доверительной вероятности  $\bar{P}$  при  $t > t(\bar{P}, f = m - 1)$ . В том случае можно говорить о наличии связи между  $X$  и  $Y$ .

Формулы для расчета границ доверительного интервала для выборочного коэффициента линейной корреляции довольно громоздки [23] и практически не используются при ручных вычислениях. Гораздо удобнее использовать номограмму, на которой по осям нанесены выборочный коэффициент корреляции  $r_{XY}$  и его истинное значение  $\rho_{XY}$  (обозначенные  $r$  и  $\rho$ ), а на кривых, каждая из которых соответствует тому или иному объему выборки  $m$ , указаны доверительные интервалы для  $\rho_{XY}$  (рис. 2.13).

С использованием той же номограммы можно проводить и графическую проверку гипотезы  $H_0: (\rho_{XY} = 0)$  об отсутствии линейной зависимости между измеряемыми величинами (вместо расчетов по формуле 2.75). Если величина  $r$  не попадает в доверительный интервал для  $\rho$ , то гипотеза отвергается.

Непараметрический коэффициент корреляции рангов (коэффициент Спирмена) используется, если величины  $X$  и  $Y$  не имеют нормального распределения. В таких ситуациях, а также при сильных отклонениях от линейной зависимости между величинами  $X$  и  $Y$  применение линейного коэффициента корреляции неэффективно и может привести к ошибочным выводам. В таких случаях оценки коэффициента корреляции можно получить, используя не сами значения  $(X_i, Y_i)$ , а их ранги. Условием этого является монотонное (необязательно линейное) изменение функции, описывающей связь  $X$  и  $Y$ . Расчеты проводятся следующим образом.

Выборки (отдельно  $X_i$  и отдельно  $Y_i$ ) ранжируют по убыванию, т. е. располагают в порядке уменьшения результатов измерений. Заменяя величины  $X_i$  их порядковыми номерами в получившемся ряду (рангами)  $A_i$ , а величины  $Y_i$  их рангами  $B_i$ , выражение для коэффициента корреляции рангов  $R$  можно записать так:

$$R = 1 - \frac{6}{m(m^2 - 1)} \sum_{i=1}^m (A_i - B_i)^2, \quad (2.76)$$

где  $m$  — число пар значений. Величина  $R$  при  $m \rightarrow \infty$  распределена асимптотически нормально с параметрами  $\mu = 0$  и  $\sigma = 1$ .

Для проверки гипотезы о независимости величин  $X$  и  $Y$  при  $4 \leq m \leq 10$  можно воспользоваться таблицей из Приложения 10. Если коэффициент  $R$ , рассчитанный по (2.76) превышает по абсолютному значению приведенное в Приложении 10 значение, гипотезу о независимости отвергают при уровне доверительной вероятности  $\bar{P}$  (двухсторонний критерий). При  $m \leq 10$  проверку гипотезы с достаточной точностью можно проводить так же, как это делается при использовании выборочного коэффициента линейной корреляции  $r_{XY}$  (см. выше).

### 2.7.2. Регрессионный анализ

Когда взаимосвязь двух величин достоверно выявлена, естественно найти вид математической зависимости между ними. Методы нахождения такой зависимости носят название *регрессионного анализа*. Здесь мы рассмотрим простейший и в то же время наиболее часто встречающийся случай линейной зависимости между двумя величинами.

Пусть, как и прежде, имеется  $m$  ( $m > 2$ ) пар значений  $(X_i, Y_i)$  измеряемых величин  $X$  и  $Y$ . По результатам проверки способами, указанными в разд. 2.7.1, гипотеза о независимости величин  $X$  и  $Y$  отвергается и предполагается линейная связь между ними. Мы будем полагать связь между величинами  $X$  и  $Y$  линейной и запишем уравнение, связывающее (на самом деле) эти величины как

$$Y = \alpha + \beta X. \quad (2.77)$$

Это — «истинное» уравнение. Как и при решении других задач математической статистики, найти точные значения коэффициентов  $\alpha$  и  $\beta$  нельзя из-за наличия случайных погрешностей и ограниченности числа измерений. Наша задача — получить уравнение

$$Y = a + bX, \quad (2.78)$$



в котором  $a$  и  $b$  — соответственно, оценки коэффициентов  $\alpha$  и  $\beta$  уравнения (2.77), причем сделать это наилучшим в конкретной ситуации способом. Это и есть задача регрессионного анализа.

Величину  $X$  называют независимой переменной (регрессором, объясняющей переменной),  $Y$  — зависимой переменной (переменной отклика), зависимость (2.77) — уравнением регрессии, величины  $\alpha$  и  $\beta$  — коэффициентами уравнения регрессии, а методы нахождения их оценок (и погрешностей этих оценок) — регрессионным анализом. Эти термины сохраняются и для многомерного случая (см. ниже, разд. 2.7.2.5). Величина  $\beta$  есть тангенс угла наклона искомой прямой; ее оценка  $b$  — тангенс угла наклона прямой, найденной в ходе регрессионного анализа. Соответственно, величина  $\alpha$  — отрезок, отсекаемый на оси ординат искомой прямой;  $a$  — отрезок, отсекаемый на оси ординат прямой, найденной при регрессионном анализе.

### 2.7.2.1. Графический и некоторые другие методы

**Графический метод.** Как и при решении других задач математической статистики, в регрессионном анализе прежде всего применяют графический метод. Если случайные погрешности малы, а диапазон, в котором лежат результаты измерений, достаточно широк, прямую, описывающую связь между  $X$  и  $Y$ , достаточно точно находят следующим способом. Наносят точки  $(X_i, Y_i)$  на координатную плоскость и проводят (обычно с помощью прозрачной линейки) прямую, относительно которой во всей области значений величин  $X$  и  $Y$  точки  $(X_i, Y_i)$  распределяются более или менее симметрично во всем диапазоне измерений (см. рис. 2.11б). В случае большого разброса результатов измерений можно проводить прямую через середины отрезков, соединяющих соседние точки, однако такой способ весьма неточен.

Недостатками графического метода являются:

- достаточно большие погрешности и некоторая субъективность нахождения прямой, особенно при значительном разбросе результатов измерений; невозможность количественно оценить погрешность (то есть найти доверительные интервалы) для найденных оценок коэффициентов  $a$  и  $b$ .

**Метод Тейла.** Простой и эффективный способ оценки величин  $a$  и  $b$  предложен Тейлом [36, 37]. Находят угловые коэффициенты прямых, которые проводятся через любую из пар точек  $(X_i, Y_i)$  и  $(X_j, Y_j)$ ,  $X_i \neq X_j$ ,  $Y_i \neq Y_j$ :

$$b_{ij} = \frac{Y_i - Y_j}{X_i - X_j} \quad (2.79)$$

и рассчитывают медиану всех значений  $b_{ij}$ . Эта и будет оценкой коэффициента  $\beta$ :  $median(b_{ij})$ . Использование медианы связано, в частности, с тем, что среди коэффициентов  $b_{ij}$  могут встречаться величины от 0 до  $\infty$  и применение среднего значения бессмысленно.

Для нахождения коэффициента  $a$  (напомним, что это оценка величины  $\alpha$  в (2.67)) обычно пользуются теоремой о том, что «правильная» прямая всегда проходит через точку  $(\bar{X}, \bar{Y})$ . Получаем:



$$a = \bar{Y} - b\bar{X}. \quad (2.80)$$

Другими словами, как в методе Тейла, так и в линейном регрессионном анализе вообще основной задачей является нахождение углового коэффициента  $b$ , после чего коэффициент  $a$  легко найти по уравнению (2.80).

Для оценки доверительного интервала для угловых коэффициентов поступают следующим образом.

– рассчитывают величину  $C_\alpha = z_{(\alpha/2)} \left[ \frac{N(N-1)(2N+5)}{18} \right]^{1/2}$ ,

где  $z_{(\alpha/2)}$  – квантиль нормального распределения. Результат расчета округляют до ближайшего целого числа;

– упорядочивают величины  $b_j$ , например, располагают по возрастанию, и приписывают каждой из них индекс  $k$ , соответствующий положению в этом упорядоченном ряду. Обозначим эти величины как  $(b^{(k)})$ ;

– находят номер  $k$ , соответствующий медиане величин  $b^{(k)}$  – это та же  $b = \text{median}(b_j)$ ;

– рассчитывают величины  $M_H = b - C_\alpha$  и  $M_B = b + C_\alpha$ .

Доверительный интервал для углового коэффициента уравнения регрессии в этом случае есть  $\{b^{(M_H)}, b^{(M_B)}\}$  при уровне доверительной вероятности  $\bar{P} = 1 - \alpha$ .

Преимуществом метода Тейла является то, что в его основе фактически не лежат предположения о распределении результатов измерений. Он также устойчив к грубым промахам. К сожалению, этот способ эффективен только при большом числе результатов измерений  $m$  (более детальных рекомендаций дать нельзя, поскольку нужно для оценки число точек зависит как от их разброса, так и от требований к точности оценок). Тем не менее можно сказать, что метод Тейла неплохо работает при  $m > 30$ .

**Метод наименьших абсолютных отклонений** [38] предполагает поиск прямой, сумма абсолютных величин вертикальных отрезков от экспериментальных точек до которой минимальна. Данный метод схож с методом наименьших квадратов Гаусса (см. ниже), однако хуже обоснован теоретически и несколько устойчивее к грубым промахам, чем МНК.

Недавно предложенный **метод «самой глубокой» регрессии** [39] предполагает поиск класса прямых, которые при вращении вокруг какой-либо точки «накрывают» максимальное число экспериментальных точек. Из совокупности прямых, удовлетворяющих этому условию, выбирается средняя. Основным преимуществом этого непараметрического метода являются устойчивость к грубым промахам.

Последние два метода пока не получили распространения в метрологии.

### 2.7.2.2. Метод наименьших квадратов Гаусса

В метрологии вообще и в метрологии химического анализа широко используется **метод наименьших квадратов в варианте Гаусса**. Он хорошо обоснован теоретически и эффективен даже при небольшом числе измерений. Естественно, эти преимущества связаны и с недостатками: в основу метода Гаусса положен целый ряд предположений, которые не всегда выполняются на практике. Кроме того, он весьма неустойчив к грубым промахам (см. ниже).

В методе наименьших квадратов Гаусса ищется прямая, удовлетворяющая следующему условию: сумма квадратов вертикальных отрезков от экспериментальных точек до искомой прямой минимальна. При этом статистический анализ проводится в рамках модели

$$Y_i = \alpha + \beta X_i + \varepsilon, \quad (2.81)$$

где  $Y_i$  — наблюдаемые значения (результаты измерений величины  $Y$ ),  $\alpha$  и  $\beta$  — математические ожидания коэффициентов уравнения регрессии,  $X_i$  — истинные значения величины  $X$ ,  $\varepsilon$  — нормально распределенная случайная погрешность  $\varepsilon \in (0, \sigma_\varepsilon^2)$ . Это означает, помимо базового предположения о линейной связи величины  $Y$  и  $X$ , выполнение следующих предположений:

- 1) погрешностями измерения величин  $X_i$  можно пренебречь, т. е.  $(s_\varepsilon/X) \ll (s_\varepsilon/Y)$  ( $s_\varepsilon$  и  $s_\varepsilon$  — стандартные отклонения измерений величин  $X$  и  $Y$  соответственно). В силу именно этого предположения в методе Гаусса минимизируется именно сумма квадратов вертикальных расстояний от экспериментальных точек до искомой прямой;
- 2) случайная погрешность измерений величин  $Y_i$ :
  - постоянна во всем диапазоне измерений и
  - имеет нормальное распределение  $N(0, \sigma_\varepsilon^2)$ .

Для получения несмещенных оценок с учетом этих предположений как раз и необходимо минимизировать сумму квадратов вертикальных отрезков от экспериментальных точек до искомой прямой. Дифференцируя выражение для суммы квадратов вертикальных отрезков и приравнявая первую производную к нулю, получаем следующие выражения для оценок  $a$  и  $b$ :

$$b = \frac{m \sum_{i=1}^m X_i Y_i - \sum_{i=1}^m X_i \sum_{i=1}^m Y_i}{m \sum_{i=1}^m X_i^2 - \left( \sum_{i=1}^m X_i \right)^2}, \quad (2.82)$$

$$a = \frac{m \sum_{i=1}^m Y_i \sum_{i=1}^m X_i^2 - \sum_{i=1}^m X_i \sum_{i=1}^m X_i Y_i}{m \sum_{i=1}^m X_i^2 - \left( \sum_{i=1}^m X_i \right)^2} = \frac{\left( \sum_{i=1}^m Y_i - b \sum_{i=1}^m X_i \right)}{m}. \quad (2.83)$$

Стандартные отклонения оцененных (рассчитанных) значений наклона  $s_b$  и отрезка, отсекаемого на оси ординат  $s_a$ , находят по формулам

$$s_b = \sqrt{\frac{m s_0^2}{m \sum_{i=1}^m X_i^2 - \left( \sum_{i=1}^m X_i \right)^2}} \quad (2.84)$$

и

$$s_a = s_b \sqrt{\frac{\sum_{i=1}^m X_i^2}{m}}, \quad (2.85)$$

где

$$s_0 = \sqrt{\frac{\sum_{i=1}^m Y_i^2 - \frac{(\sum_{i=1}^m Y_i)^2}{m} - \left( \sum_{i=1}^m X_i Y_i - \frac{\sum_{i=1}^m X_i \sum_{i=1}^m Y_i}{m} \right)}{\sum_{i=1}^m X_i^2 - \frac{(\sum_{i=1}^m X_i)^2}{m}}} \quad (2.86)$$

есть стандартное отклонение разностей между экспериментальными величинами  $Y_i$  и найденной зависимостью (по оси ординат), т. е. между точками с координатами  $(X_i, Y_i)$  и  $(X_i, \hat{Y}_i = a + bX_i)$ . Эту величину называют также остаточной суммой отклонений, поскольку ее можно рассматривать как стандартное отклонение так называемых остатков  $e_i$  — величин, которые получаются после вычитания из экспериментальных величин  $Y_i$  соответствующих им ординат уравнения регрессии (см. ниже).

Доверительные интервалы для  $b$  и  $a$  рассчитывают как

$$\Delta b = b \pm t(\bar{P}, f) s_b \text{ и } \Delta a = a \pm t(\bar{P}, f) s_a, \quad (2.87)$$

где  $t(\bar{P}, f)$  есть процентная точка  $t$ -распределения Стьюдента, соответствующая доверительной вероятности  $\bar{P}$  и числу степеней свободы  $f = m - 1$ .

Полученное уравнение регрессии  $Y = a + bX$  может быть использовано для расчета так называемых «предсказанных» значений  $Y_k$  по заданным значениям  $X_k$ . Другими словами, предсказанные значения  $Y_k = a + bX_k$  могут быть рассчитаны для любого  $X_k$  после того, как методом Гаусса найдено уравнение регрессии (2.78).

Представляет интерес оценка погрешности величин  $Y_k$ . Доверительный интервал для каждой  $Y_k$  можно оценить по формуле

$$\Delta Y_k = Y_k \pm t(\bar{P}, f) s_0 \sqrt{\frac{1}{m} + \frac{m(X_k - \bar{X})^2}{m \sum_{i=1}^m X_i^2 - (\sum_{i=1}^m X_i)^2}} \quad (2.88)$$

Здесь, как и выше,  $t(\bar{P}, f)$  есть процентная точка  $t$ -распределения Стьюдента, соответствующая доверительной вероятности  $\bar{P}$  и числу степеней свободы  $f = m - 2$ .

Если мы найдем границы доверительных интервалов для всех возможных  $Y_k$ , то получившиеся линии определяют границы, в которых лежат все уравнения регрессии, возможные при доверительной вероятности  $\bar{P}$  (рис. 2.14). Как видно, доверительные интервалы для  $Y_k$  быстро увеличиваются с удалением от средней точки  $(\bar{X}, \bar{Y})$ . Особенно большие погрешности возникают при выходе за пределы диапазона, в котором лежат  $X_i$ .

Иногда пытаются рассчитать коэффициенты  $a$  и  $b$  по формулам (2.82) и (2.83), поменяв местами  $X$  и  $Y$ . То, что при этом получается совершенно другая линия регрессии, у многих вызывает недоумение. Однако это совершенно естественно: данные формулы справедливы лишь для модели (2.81) при выполнении соответствующих предположений (см. выше); использование их после замены  $X$  на  $Y$  и наоборот (то есть когда независимая переменная оказывается в левой части уравнения регрессии) некорректно.



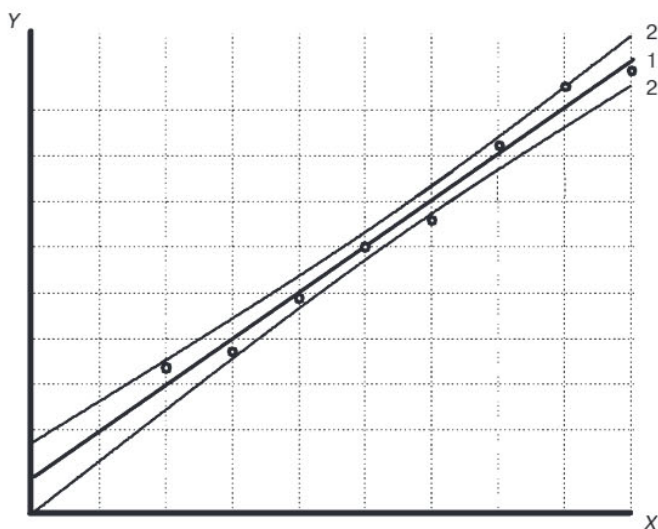


Рис. 2.14. Доверительные интервалы для уравнения регрессии. 1 — график найденного уравнения регрессии  $Y = a + bX$ , 2 — границы доверительных интервалов

### 2.7.2.3. Случаи невыполнения основных предположений

Метод наименьших квадратов Гаусса хорошо работает только в случае выполнения указанных в разд. 2.7.2.2 предположений. Однако на практике они выполняются не всегда. Рассмотрим наиболее частые ситуации.

**Ситуация 1.** Погрешностями измерения величины  $X$  нельзя пренебречь, то есть величины  $s_X/X$  и  $s_Y/Y$  ( $s_X$  и  $s_Y$  — стандартные отклонения измерений величин  $X$  и  $Y$  соответственно) соизмеримы. Такая ситуация возникает, например, при сравнении двух методов анализа путем исследования обоими методами набора проб с разной концентрацией определяемого вещества, либо при градуировке по стандартным образцам, аттестованным со значительной погрешностью (гл. 3). При этом результаты измерений можно описать моделью

$$Y_i = \alpha + \beta(X_i + \xi) + \varepsilon. \quad (2.89)$$

Другими словами, предполагается, что между величинами  $X$  и  $Y$  на самом деле имеется линейная зависимость  $Y = \alpha + \beta X$ , но каждая из них имеет собственную погрешность:

$$Y_i = Y_i + \varepsilon \text{ и } X_i = X_i + \xi. \quad (2.90)$$

Здесь  $Y_i$  и  $X_i$  — истинные значения (математические ожидания совокупностей) результатов измерений величин  $Y$  и  $X$ . Обычно полагают, что случайные погрешности  $\xi$  и  $\varepsilon$  имеют нормальное распределение:  $\xi \in N(0, \sigma_\xi^2)$ ,  $\varepsilon \in N(0, \sigma_\varepsilon^2)$ .

В этом случае прямая, некорректно найденная по методу Гаусса с использованием формул (2.82, 2.83), оказывается повернутой относительно истинной прямой по часовой стрелке вокруг точки  $(\bar{X}, \bar{Y})$  (рис. 2.15).

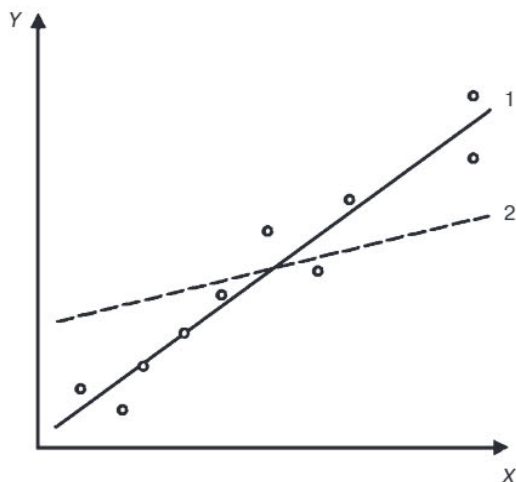


Рис. 2.15. Последствия неучета погрешности измерения независимой переменной. 1 — истинная прямая; 2 — прямая, ошибочно найденная методом Гаусса без учета погрешности измерения величины  $X$

Если дисперсии величин  $X$  и  $Y$  ( $\sigma_x^2$  и  $\sigma_y^2$ ) постоянны во всем диапазоне измеряемых величин и известны их оценки  $s_x^2$  и  $s_y^2$ , коэффициенты «правильного» уравнения регрессии можно рассчитать следующим способом [40]:

- найти (неправильные) оценки коэффициентов в уравнении регрессии  $a$  и  $b$  методом Гаусса по формулам (2.72) и (2.73);
- рассчитать скорректированный коэффициент  $b^*$  по формуле

$$b^* = \frac{1}{2} \left[ \frac{b}{r^2} - \frac{\lambda}{b} + \sqrt{\left( \frac{b}{r^2} - \frac{\lambda}{b} \right)^2 + 4\lambda} \right], \quad (2.91)$$

где  $r$  — выборочный коэффициент корреляции между величинами  $X$  и  $Y$  (см. разд. 2.7.1),  $\lambda = s_y^2/s_x^2$ ;

- рассчитать скорректированный коэффициент  $a^*$  из условия, что искомая прямая проходит через точку  $(\bar{X}, \bar{Y})$ :

$$a^* = \bar{Y} - b^* \bar{X}. \quad (2.92)$$

Искомая прямая  $Y = a^* + b^*X$ .

Полезно знать, что оценки  $b^*$ , полученные по формуле (2.91), имеют следующие свойства:

- $b^*$  всегда больше  $b$ ;
- $b^*$  стремится к  $b$  при увеличении  $\lambda$ ;
- $b^*$  стремится к  $b/r^2$  при малых  $\lambda$ ;
- разность между  $b^*$  и  $b$  растет при уменьшении коэффициента корреляции  $r$ ;
- разность между  $b^*$  и  $b$  наиболее чувствительна к изменению величины  $\lambda$  при  $\lambda \approx 1$ .



Другой способ нахождения линейной зависимости при значительной погрешности регрессора  $X$  – применение метода главных компонент (см., например, [41]).

**Ситуация 2.** Погрешностями измерения величин  $X_i$  можно пренебречь, однако случайная погрешность измерений величин  $Y_i$  существенно зависит от  $X$  («гетероскедастичность»). В практике количественного химического анализа вещества такая ситуация встречается очень часто, особенно если диапазон измеряемых концентраций достаточно широк.

Обычным способом действия в такой ситуации является применение «взвешенного» метода наименьших квадратов. Рассмотрение ведется в рамках модели, аналогичной (2.71):  $Y_i = \alpha + \beta X_i + \varepsilon$ , однако случайная погрешность  $\varepsilon$  полагается нормально распределенной с параметрами  $0, \sigma_\varepsilon^2 w_i$ ;  $\varepsilon \in N(0, \sigma_\varepsilon^2 w_i)$ , где  $w_i$  – «веса», то есть положительные числа, показывающие, во сколько раз дисперсия при  $i$ -м измерении меньше принятой за стандартную дисперсии  $\sigma_\varepsilon^2$ . Например, при количественном химическом анализа величины  $w_i$  часто бывают приблизительно пропорциональны  $X^2$ , то есть отношение  $\sigma_\varepsilon^2/X$  практически постоянно.

При использовании «взвешенного» метода наименьших квадратов несмещенные оценки коэффициентов  $a$  и  $b$  получаются путем минимизации суммы взвешенных, то есть умноженных на соответствующие веса  $w_i$ , квадратов вертикальных отрезков от экспериментальных точек до искомой прямой.

Обозначив  $\bar{Y}_w = \frac{\sum_{i=1}^m w_i Y_i}{\sum_{i=1}^m w_i}$ ,  $\bar{X}_w = \frac{\sum_{i=1}^m w_i X_i}{\sum_{i=1}^m w_i}$ , для взвешенного метода наименьших

квадратов получаем следующие расчетные формулы для коэффициентов уравнения регрессии  $Y = a + bX$ :

$$b = \frac{\sum_{i=1}^m w_i (Y_i - \bar{Y}_w)(X_i - \bar{X}_w)}{\sum_{i=1}^m w_i (X_i - \bar{X}_w)^2}, \quad a = \bar{Y}_w - b\bar{X}_w. \quad (2.93)$$

Как показывает опыт, точного знания весов  $w_i$  не требуется. Если выявлена существенная гетероскедастичность (например, с помощью диаграммы остатков, см. разд. 2.7.2.4), обычно экспериментально оценивают зависимость  $\sigma_\varepsilon^2$  от  $X$ . Для этого проводят по несколько измерений для нескольких значений  $X$  и для каждого рассчитывают дисперсию  $s_\varepsilon^2$ . Исходя из найденной зависимости оценивают  $w_i$ . Известны и более точные итерационные процедуры (см. [42]).

**Ситуация 3.** Возможны более сложные случаи – сильная гетероскедастичность при наличии значительной погрешности измерения независимой переменной. Корректные алгоритмы действий в таких ситуациях достаточно сложны и не всегда устойчивы [42–44]. К счастью, при решении метрологических задач методов, описанных выше для ситуаций 1 и 2, достаточно для получения достаточно точных результатов.

**Ситуация 4.** Наконец, бывает, что зависимость  $Y = f(X)$  нелинейна. Это может быть известно заранее из теории данного метода измерений.

**Пример.** Зависимость концентрации определяемого вещества  $C$  от времени  $t$  при исследовании кинетики реакций первого порядка

$$C = C_0 \exp(-kt), \quad (2.94)$$

где  $C_0$  — константа, равная начальной концентрации,  $k$  — константа скорости реакции.

Также нелинейность часто выявляется по результатам графического рассматривания.

В этом случае можно действовать двумя способами.

- Применить методы нелинейной регрессии (см., например, [45, 46]). Многие алгоритмы нелинейной регрессии реализованы в популярных компьютерных программах для статистических расчетов. К сожалению, зачастую они оказываются неустойчивыми в конкретных ситуациях.
- Использовать линеаризующее преобразование (функцию, обратную  $f(X)$ , то есть  $Z = f^{-1}(X)$ ) и пересчитать данные по формуле

$$Z_i = f^{-1}(Y_i). \quad (2.95)$$

При этом мы перейдем в координаты, в которых зависимость новой переменной  $Z$  от  $X$  линейна<sup>34</sup>. Так, для вышеприведенного примера (2.94) линеаризующее преобразование  $Z = \ln C$ . Далее можно использовать метод наименьших квадратов Гаусса и другие методы, при необходимости учитывая гетероскедастичность, которая может возникнуть из-за проведенного преобразования. После нахождения линейной зависимости в преобразованных координатах возвращаются к исходным координатам.

Оба подхода корректны, и выбор определяется наличием у исследователя того или иного программного обеспечения, а также особенностями решаемой задачи.

#### 2.7.2.4. Анализ остатков. Проверка выполнения предположений и выявление грубых промахов

Невыполнение предположений, лежащих в основе метода наименьших квадратов Гаусса, сильно влияет на результаты расчетов. Рекомендуется всегда проверять, выполняются ли эти предположения, и при необходимости использовать методы, рассмотренные в разделе 2.7.2.3. Одним из наиболее эффективных способов такой проверки является *анализ остатков*.

Напомним, что *остатками* называют величины  $e_i = [Y_i - \hat{Y}_i]$ , где  $\hat{Y}_i = a + bX_i$  ( $a$  и  $b$  — оценки коэффициентов уравнения линейной регрессии, найденные методом Гаусса), см. разд. 7.2.2). Другими словами, это вертикальные отрезки от экспериментальных точек до прямой, рассчитанной методом Гаусса. Рассматривая остатки в графической форме, можно получить обширную информацию (рис. 2.16).

При выполнении основных предположений (то есть в отсутствие гетероскедастичности, малой погрешности измерений величин  $X_i$  и при линейном виде зависимости) и при корректно проведенных расчетах точки  $e_i = [Y_i - \hat{Y}_i]$  распределены вокруг оси абсцисс случайным образом (рис. 2.16а). Если дисперсия результатов измерений  $Y_i$  неоднородна, то разброс точек меняется с изменением величины  $X$

<sup>34</sup> Строго говоря, такое преобразование не всегда возможно. Однако при решении метрологических задач линеаризующее преобразование почти всегда удастся подобрать с приемлемой погрешностью.

(пример с обычно встречающимся ростом случайной погрешности при увеличении  $X$  приведен на рис. 2.16б). На графике остатков очень четко обнаруживаются отклонения от линейной зависимости (на рис. 2.16в показан случай, когда зависимость между  $Y$  и  $X$  на самом деле представляет собой часть синусоиды). И, наконец, на рис. 2.16г показан случай некорректного пренебрежения погрешностями измерений величин  $X$ . Остатки распределяются вокруг прямой, проходящей через точку  $(\bar{X}, 0)$  и повернутой относительно оси  $X$  по часовой стрелке; эта прямая показана пунктиром.

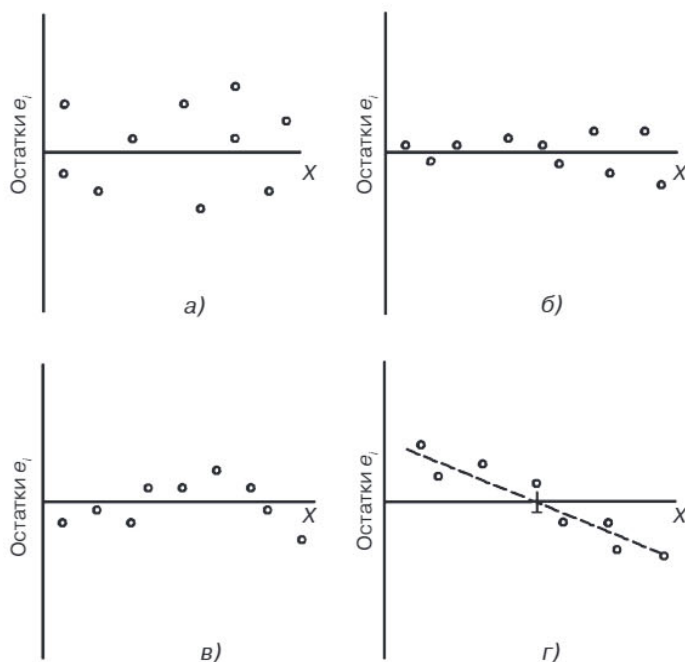


Рис. 2.16. Графическое изображение остатков: основные предположения выполняются (а); заметная гетероскедастичность (разброс результатов измерений растет с ростом  $X$ ) (б); нелинейная зависимость (в); не учтена существенная погрешность измерения независимой переменной (г)

#### *Выявление грубых промахов при использовании метода Гаусса.*

Возможны также количественные расчеты с использованием остатков, важнейшим из которых является выявление грубых промахов.

В разд. 2.4.3 обсуждалось, как выявлять грубые промахи при повторных измерениях одной и той же величины (одномерный случай). Также проблема возникает и при регрессионном анализе: выбросы могут катастрофически исказить искомую зависимость. Рассмотрим кратко некоторые ее аспекты для двумерного случая поиска линейной зависимости ( $Y = a + bX$ ).

В связи с проблемой подозрительно выделяющихся значений и грубых промахов в регрессионном анализе полезным понятием является «точка неустойчивости» (вольный перевод с англ. «breakdown point», BDP). Она представляет собой процент



подозрительно выделяющихся значений, которые могут быть отброшены без риска, что регрессионная линия значительно изменится. Можно показать, что эта величина не может превышать 50%. Особый интерес вызывают методы с  $BDP > 20\%$ . Таковы описанный выше алгоритмы Тейла и метод «самой глубокой» регрессии. *В то же время для метода наименьших квадратов BDP равна нулю, поскольку даже отброс одной точки может сильно изменить уравнение регрессии.*

Таким образом, проблема подозрительно выделяющихся значений и грубых промахов стоит не слишком остро, если для нахождения линейной зависимости применяется графический метод, метод Тейла либо метод «самой глубокой» регрессии. Действительно, при использовании графического метода одну явно выпадающую точку можно проигнорировать, проводя прямую только по остальным точкам. Метод Тейла устойчив к выбросам, поскольку в нем используется медиана — параметр распределения, на который не влияет малое число выпадающих значений. Не реагирует на некоторое количество грубых промахов и метод «самой глубокой» регрессии. Однако *даже один грубый промах может сильно исказить уравнение регрессии при использовании метода наименьших квадратов.*

Известно много способов выявления грубых промахов в регрессионном анализе. Большинство относительно простых из них основано на анализе остатков; именно они включены в наиболее известные пакеты статистических программ.

Простейший способ заключается в том, что остатки  $e_i = [Y_i - \hat{Y}_i]$  рассматриваются как одномерная выборка. Если среди них остатков есть подозрительно выделяющееся значение, казалось бы, можно проверить, является ли оно грубым промахом, используя методы, описанные в разд. 2.4.3. Однако применение параметрических методов непосредственно к величинам  $e_i$  некорректно, поскольку они не являются независимыми, и может привести к ошибкам (особенно при небольшом числе измерений). Взаимно независимыми оказываются «*студентизованные*» *остатки*:

$$r_i = \frac{e_i}{s_0(1-h_i)^{1/2}}, \quad (2.96)$$

где  $s_0$  — остаточная сумма отклонений, рассчитываемая по формуле (2.86), а

$$h_i = \frac{1}{m} + \frac{(X_i - \bar{X})^2}{\sum_{i=1}^m (X_i - \bar{X})^2}. \quad (2.97)$$

К величинам  $r_i$  применимы все методы выявления грубых ошибок, описанные в разд. 2.4.3. Заметим, что, поскольку величины  $h_i$  характеризуют расстояние от  $X_i$  до среднего, их используют для дополнительного выявления грубых ошибок в случае, когда погрешностями измерений  $X_i$  нельзя пренебречь.

Иногда вместо  $r_i$  используют «*джекнайф*» *остатки*:

$$r_{-i,i} = \frac{e_i}{s_i(1-h_i)^{1/2}}, \quad (2.98)$$

где  $s_i$  — аналог величины  $s_0$  с той разницей, что  $i$ -я точка исключена из расчетов.

При большом числе измерений «студентизованные» и «джекнайф» остатки мало отличаются.

Еще один способ выявления выбросов в регрессионном анализе основан на использовании *расстояния Кука*  $d_i$  – статистики, показывающей степень изменения коэффициентов уравнения регрессии при исключении отдельной точки

$$d_i = \frac{e_i h_i}{s_i (1 - h_i)^{1/2}}. \quad (2.99)$$

Можно показать, что величины  $d_i$  всегда больше нуля и в норме не превышают 1. Если  $d_i > 1$ , то соответствующую точку рассматривают как выброс. Подробнее см. [47].

### 2.7.2.5. Многомерная регрессия

Предположим, что в отличие от рассмотренного выше одномерного случая случайная величина  $Y$  зависит не от одной величины  $X$ , а от нескольких величин  $Z_1, Z_2, \dots, Z_k$  и для каждого экспериментального определения величины  $Y$  известны значения каждой из величин  $Z$ . Другими словами, имеем набор данных  $(Y_i, Z_{1i}, Z_{2i}, \dots, Z_{ki})$ ,  $i = 1, 2, \dots, m$ , соответствующий  $m$  точкам в многомерном (размерностью  $k + 1$ )-пространстве. Необходимо найти зависимость  $Y$  от величин  $Z_1, Z_2, \dots, Z_k$ . В предположении о линейном характере связи величины  $Y$  с каждой из независимых переменных  $Z_1, Z_2, \dots, Z_k$  результаты обычно обрабатывают в рамках модели<sup>35</sup>

$$Y_i = \alpha + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \dots + \beta_k Z_{ki} + \epsilon, \quad (2.100)$$

где  $Z_{1i}, Z_{2i}, \dots, Z_{ki}$  – математические ожидания величин  $Z_1, Z_2, \dots, Z_k$ ,  $\epsilon$  – нормально распределенная случайная погрешность  $\epsilon \in N(0, \sigma_\epsilon^2)$ . Требуется найти оценки коэффициентов  $\alpha, \beta_1, \beta_2, \dots, \beta_k$ , характеризующие связь зависимой переменной  $Y$  с независимыми переменными  $Z_1, Z_2, \dots, Z_k$ . Такая модель – обобщение модели (2.81) для случая нескольких независимых переменных. Фактически все сводится к поиску прямой в  $(k + 1)$ -мерном пространстве.

В химическом анализе такая задача решается, например, при изучении зависимости результатов определения концентрации какого-либо анализата от концентрации других веществ, присутствующих в анализируемой пробе.

Чаще всего используют метод наименьших квадратов Гаусса в многомерном варианте, в котором в многомерном пространстве ищется прямая, сумма квадратов расстояний от экспериментальных точек до которой минимальна. Основные предположения стандартны для метода Гаусса: отсутствие (малость) погрешностей измерения величин  $Z_1, Z_2, \dots, Z_k$ , отсутствие гетероскедастичности, нормальное распределение погрешности измерений. Однако в многомерном варианте возникает дополнительное ограничение: величины  $Z_1, Z_2, \dots, Z_k$  должны быть взаимно независимыми. Так, при поиске связи выхода продукта с условиями проведения

<sup>35</sup> Иногда рассматривают также взаимодействие переменных  $Z_1, Z_2, \dots, Z_k$  между собой. Однако учет таких эффектов требует большого объема экспериментальных данных и в метрологии химического анализа встречается редко.



реакции в замкнутом объеме нельзя включать в модель (2.00) давление и температуру одновременно — результаты будут неправильными и непредсказуемыми.

Методы решения задачи многомерной регрессии с использованием метода Гаусса описаны в [42, 44, 47, 48]. Если погрешностями измерения величин  $X_k$  нельзя пренебречь, оценки коэффициентов  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ , ...,  $\beta_k$  можно найти методом главных компонент [48]. Возможность проведения подобных расчетов заложена в большинство имеющихся на рынке компьютерных программ, предназначенных для статистической обработки данных.

## ГЛАВА 3

# МЕТРОЛОГИЯ ХИМИЧЕСКОГО АНАЛИЗА

Как было указано в гл. 1, построить для химического анализа классическую систему передачи размера от единого первичного эталона к образцовым и рабочим средствам измерений практически невозможно. Постоянно возникают даже дискуссии о том, является ли химический анализ измерением — столь велика его специфика. В настоящее время с учетом особенностей количественного химического анализа его выделяют в специфический вид измерений, которым занимается *метрология химического анализа*. Учитывая принципиальное значение вопроса о специфике КХА для настоящей книги, обсудим его подробнее.

### 3.1. Основные принципы метрологии химического анализа

*Метрология химического анализа* — раздел науки, лежащий на стыке *метрологии* и *аналитической химии*, в котором определение содержания компонентов анализируемого вещества — количественный химический анализ (КХА) рассматривается как специфический вид измерений.

Почти всегда результат КХА находят с помощью известной теоретически либо (чаще) экспериментальной зависимостью между концентрацией определяемого вещества (*аналита*) и непосредственно измеряемыми величинами или их комбинациями (*аналитический сигнал*). Такими величинами могут быть интенсивность спектральных линий при атомно-эмиссионном анализе, количество реагента, вступившего в реакцию с аналитом при объемном анализе; количество электричества, затраченное на окисление либо восстановление аналита при кулонометрии, и т. д.

Следовательно, КХА почти всегда является *косвенным измерением* (см. гл. 1). При этом от других («классических») видов измерений КХА отличается [49]:

1. Отсутствием единого эталона единицы количества вещества — моля.
2. Сложностью и многостадийностью методик КХА, включающих в большинстве случаев пробоподготовку (измельчение пробы, ее растворение и т. д.), концентрирование определяемых веществ (экстракция, сорбция) и другие операции. *При этом одни и те же вещества (аналиты) в разных объектах определяются по разным методикам.* Все это, в свою очередь, определяется спецификой решаемой задачи — необходимо получить аналитический сигнал, связанный с количеством одного или нескольких веществ, присутствующих в веществе сложного состава, часто при низкой их концентрации.
3. Характерным для многих методов химического анализа наличием значительных по величине скрытых систематических погрешностей, обусловленных матричными эффектами (влиянием состава конкретного образца на результаты анализа).

Рассмотрим, как эти особенности определяют современную метрологию химического анализа.

### *3.1.1. Возможна ли корректная передача размера в количественном химическом анализе?*

При постановке проблемы построения для количественного химического анализа схемы передачи размера, подобной показанной на рис. 1.1, необходимо ответить на следующие основные вопросы: что может быть эталоном в анализе вещества? что представляют собой образцовые меры в химическом анализе? что есть средства измерений в химическом анализе?

#### *Что может быть первичным эталоном в КХА?*

Приведем следующие соображения. Как известно, за единицу количества вещества в системе СИ и принят моль — количество вещества, содержащее  $6,022140857 \times 10^{23}$  структурных единиц (атомов, молекул ионов, свободных радикалов и т.д.) (см. Приложение 1). Из этого определения ясно, что единый эталон моля создать очень сложно — ведь в определении задано лишь число структурных элементов (при количественном химическом анализе — обычно атомов или молекул) в одном моле. Тем не менее выход есть, и он используется. Для воспроизведения размера моля какого-либо вещества (в качестве эталона) может служить навеска образца этого вещества достаточно высокой чистоты<sup>36</sup>.

Предлагалось [50] и альтернативное решение — использовать в качестве первичного эталона моля чистый образец кристаллического йода I<sub>2</sub> и организовать систему передачи размера моля к другим химическим элементам с использованием их свойства образовывать простые стехиометрические соединения с йодом. Такая система теоретически возможна, но на практике неосуществима — накапливающиеся при многократных операциях передачи размера погрешности быстро превысят допустимые. Более того, в ней и нет необходимости.

Современные методы исследований (в частности, масс-спектрометрия) позволяют устанавливать относительную атомную (молекулярную) массу с точностью, значительно превышающей практические потребности. Формулы химических соединений чаще всего хорошо известны, поэтому при достаточно высокой чистоте индивидуального вещества размер моля оказывается воспроизведенным при взятии навески.

Большинство методов химического анализа имеет общую погрешность не менее 2–3%. Следуя принципу пренебрежения малыми погрешностями (см. разд. 2.6.1), можно считать, что, если суммарная массовая доля примесей в каком-либо веществе не превышает 0,7–1%, то ею можно пренебречь. Используя современные методы очистки веществ, можно получить такие и даже гораздо более чистые образцы и доказать их чистоту, например, оценив верхнюю границу суммарной концентрации возможных примесей (о стандартных образцах состава чистых веществ см. разд. 3.2). Колебаниями в изотопном составе веществ природного происхождения, очевидно, также можно пренебречь, а при определении содержания изотопов в качестве эталонов можно использовать соединения этих изотопов либо вносить соответствующие поправки.

<sup>36</sup> Конечно, при этом вещество должно быть стабильным. Для нестабильных веществ приходится искать обходные пути.



Далее. Система передачи размера массы хорошо развита, и на этом пути трудностей нет. Всегда можно выбрать весы, погрешностью взвешивания на которых при воспроизведении размера моля можно пренебречь. Действительно, используя обычные аналитические лабораторные весы с абсолютной погрешностью  $10^{-4}$  г при массе навески 0,1–1 г, имеем относительную погрешность 0,01–0,1%, заведомо несущественную по сравнению с погрешностью самых точных методов количественного химического анализа.

*Таким образом, имея достаточно чистое индивидуальное вещество (например, нитрат кальция или ацетальдегид) и зная, как это чаще всего бывает, его химическую формулу, мы можем воспроизвести моль этого вещества, взяв его навеску на достаточно точных весах.*

В случаях, когда формула соединения неизвестна, корректное воспроизведение размера моля взвешиванием чистого вещества невозможна и приходится изыскивать паллиативные решения. Чаще всего получаемый некоторым определенным способом образец по общему согласию используется в качестве первичного эталона.

**Пример.** Так обстоит дело со многими белками. Хотя аминокислотная последовательность многих из них известна, важнейшие свойства, собственно и делающие их белками, проявляются только тогда, когда в структуре присутствует некоторое количество воды. Поэтому в качестве эталона берут образец, очищенный и умеренно высушенный по некоторой общепринятой процедуре [51].

#### *Что представляют собой образцовые меры в химическом анализе?*

Из сказанного выше понятно, что роль первичного эталона (меры высшего уровня) в КХА может играть навеска чистого вещества. Если бы можно было этим ограничиться, то для каждого аналита мы должны были бы иметь один эталон — чистое вещество. К сожалению, этого недостаточно.

В качестве рабочей меры низшего уровня приходится использовать навески веществ, близкие по составу и свойствам к анализируемым пробам: горные породы, сплавы, водные растворы, высушенные продукты питания и т. д. Это обусловлено следующими причинами:

- концентрацию одних и те же аналитов в разных объектах определяют разными методами. Это связано с разным составом и свойствами анализируемых веществ. При этом один и тот же аналит в одном объекте исследований может определяться разными методами;
- большинство применяемых на практике методов анализа имеют скрытые систематические погрешности, зависящие от состава и свойств анализируемых объектов (см. ниже).

При этом в развитой иерархии эталонов, подобной показанной на рис. 1.1, чаще всего нет необходимости, поскольку погрешности, с которыми известен состав чистых веществ, ненамного меньше погрешностей рутинного химического анализа.

Таким образом, в большинстве случаев для построения метрологии КХА достаточно двух видов мер: навески чистого вещества и вещества, близкого по составу и свойства к анализируемым пробам. *Вещества и материалы*<sup>37</sup>, используемые

<sup>37</sup> Под материалами мы подразумеваем вещества, подвергнутые специальной обработке (измельчение, растворение и т. д.).



в химическом анализе в метрологических целях, носят название *образцов сравнения и стандартных образцов* (см. ниже, разд. 3.2).

*Что есть средства измерений в химическом анализе?*

Средства измерений в классическом смысле (см. гл. 1), такие как весы, вольтметры и т. д., широко используются при КХА. Они применяются для измерения величин и на промежуточных стадиях анализа (измерение объема, массы при отборе пробы и дозировании химических реактивов), и на конечной — при измерении аналитического сигнала, непосредственно связанного с концентрацией определяемого вещества. Однако эти средства измерений для большинства методик КХА не вносят определяющего вклада в общую погрешность анализа: чаще всего их погрешности намного меньше общей погрешности химического анализа.

*Отсюда следует, что при обеспечении точности количественного химического анализа следует ориентироваться не столько на точность средств измерений, сколько на точность методики в целом*<sup>38</sup>.

Очевидно, для передачи размера моля от стандартного образца чистого вещества (первичный стандартный образец согласно [52]) к рабочей мере (стандартному образцу сложного состава, вторичному стандартному образцу согласно [52]) необходима образцовая методика выполнения измерений, общая погрешность которой была бы мала по сравнению с требуемой погрешностью аттестации этого образца.

С учетом всего сказанного можно построить схему передачи размера для количественного химического анализа (рис. 3.1).

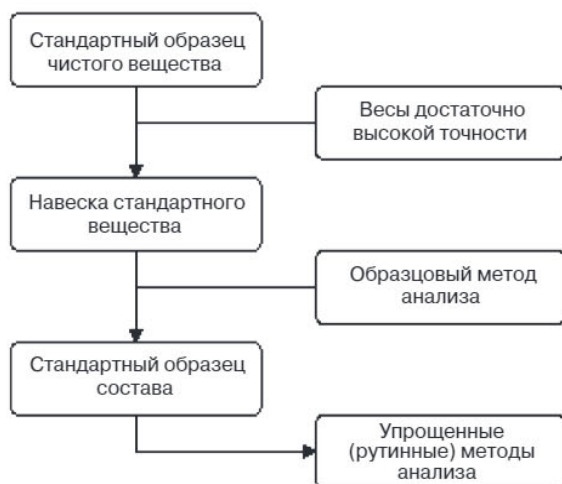


Рис. 3.1. Схема воспроизведения и передачи размера моля в отсутствие скрытых систематических погрешностей

Такая схема действительно реализована для некоторых видов КХА. В качестве примера приведем схему передачи размера моля при определении содержания холестерина в сыворотке крови человека (Национальный институт стандартов и

<sup>38</sup> Это не относится к простым автоматическим анализаторам, таким как газоанализаторы, реализующим все этапы аналитического процесса: от отбора проб до выдачи результата.

технологий, США). В качестве эталона моля холестерина используется навеска стандартного образца холестерина #909 с содержанием основного вещества более 99,7%. На более низком уровне находится стандартный образец #914, представляющий собой специальным образом стабилизированную и расфасованную сыворотку крови человека. Для передачи размера к стандартному образцу #914 в качестве средства измерения используется очень точный (общая относительная погрешность менее 1% [53]) метод определения холестерина с помощью масс-спектрометрии с изотопным разбавлением (см. ниже, п. 3.1.4), при этом градуировку осуществляют по стандартному образцу #909. Аттестованный таким способом стандартный образец #914 используется для оценки погрешностей более простых методов определения холестерина<sup>39</sup>.

*Такая схема, вполне корректная с точки зрения классической метрологии, редко используется на практике.* Дело в том, что в подавляющем большинстве случаев образцовый метод анализа для передачи размера моля стандартным образцам сложного состава с требуемой точностью отсутствует. Если же такой метод существует, то, как правило, его использование обходится очень дорого. Это наглядно видно на приведенном выше примере: одно измерение методом масс-спектрометрии с изотопным разбавлением стоит несколько тысяч долларов США. Более того, *передача размера по классической схеме зачастую вообще не имеет смысла, поскольку рутинным методам анализа присущи скрытые систематические погрешности, искажающие результаты анализа конкретного стандартного образца.*

Следовательно, передача размера по классической схеме, от единого первичного эталона к образцовым и рабочим средствам измерений (см. гл. 1) в химическом анализе во многих случаях, особенно для объектов сложного состава, невозможна. Метрологические схемы приходится выстраивать по-разному для разных аналитов и методик КХА с учетом их специфики с использованием образцов сравнения и стандартных образцов (см. также ниже, разд. 3.2).

### **3.1.2. Особенности методик химического анализа как разновидности методик измерений**

Таким образом, количественный химический анализ представляет собой весьма специфический вид измерений. Для обеспечения приемлемых погрешностей методик КХА используются методы и приемы, не применяемые вовсе или применяемые крайне редко при других видах измерений.

Как указано в разд. 3.1.1, *отсутствие эталона моля приводит* к необходимости в качестве мер использовать стандартные образцы состава веществ.

*Сложность и многостадийность* большинства методик количественного химического анализа состоит в том, что каждую из них можно рассмотреть как последовательность более или менее независимых этапов — отбор пробы, подготовка ее к анализу (в том числе выделение и концентрирование аналита), собственно измерение, обработка результатов — каждый из которых во многих случаях состоит из более мелких стадий.

<sup>39</sup> Подробнее о методах определения холестерина в сыворотке крови см. [53].



Все это — следствие сложности получения аналитического сигнала, связанного с концентрацией одного или нескольких веществ, входящих в состав объектов сложного состава (да еще во многих случаях в очень низкой концентрации).

Погрешности каждого из этих этапов вносят вклад в общую погрешность результатов анализа. При этом сложность большинства методик химического анализа как процесса измерений определяется как их многостадийностью, так и спецификой выполняемых операций. Как правило, на каждом этапе проводится совокупность операций (отбор пробы, градуировка, перенос вещества из одной емкости в другую, перемешивание, дозирование или взвешивание и т.д.), каждая из которых выполняется вручную и чревата как систематическими и обычными случайными погрешностями, так и грубыми промахами. Известно, что погрешности анализа определяются как качеством используемой аппаратуры и реактивами, состоянием окружающей среды, так и квалификацией и даже состоянием здоровья и настроением аналитика.

В результате оказывается, что даже простое воспроизведение известной, хорошо отработанной методики КХА есть достаточно сложная задача. То же касается и обеспечения качества уже реализованной методики — поддержания ее в статистически контролируемом состоянии, исключения влияния на результаты измерений грубых промахов и т. д. (эти вопросы рассмотрены ниже, см. гл. 4).

Одна из основных и трудноразрешимых метрологических проблем анализа вещества — *скрытые систематические погрешности*<sup>40</sup>. Эти погрешности имеют место в абсолютном большинстве методик КХА из-за влияния состава и физико-химических свойств проб на результаты анализа. Скрытыми эти погрешности называют потому, что, поскольку измерение концентрации всех мешающих компонентов или физико-химических характеристик при каждом измерении, как правило, не проводится, узнать об их наличии в ходе самого измерения нельзя. Для иллюстрации проблемы приведем пример, составленный по данным нашей работы [54].

**Пример.** Для выбора состава образцов сравнения для биохимических исследований изучалось влияние матрицы на результаты определения общего (суммы свободного и этерифицированного) холестерина в сыворотке крови человека, для чего использовали распространенные в рутинных анализах фотометрические методы: (А) — прямой химический с реактивом Либермана — Бурхарда; (В) — экстракционный химический (автоанализатор АА-II «Technicon»); (С) и (D) — ферментные методы, отличающиеся реагентами и способом регистрации поглощения.

Было исследовано около 400 проб сыворотки крови человека, в каждой из которых определяли содержание общего холестерина указанными методами и методом Абея, имеющим намного меньшие, чем рутинные методы, погрешности и выбранным в качестве метода сравнения. Кроме того, в каждой пробе измеряли содержание других веществ, которые потенциально могли влиять на результаты определения общего холестерина (мешающие компоненты).

По результатам измерений для каждой пробы и по отношению к каждому рутинному методу анализа была рассчитана разность *B* между результатами рутинного метода и метода сравнения (смещение). Фактически это — оценка полной систематической погрешности. Далее методом многомерного регрессионного анализа (см. разд. 2.7.2.5) была найдена за-

<sup>40</sup> Иногда такие погрешности называют «не выявленными», подразумевая, что их можно выявить и учесть, внося соответствующие поправки. Конечно, в принципе это возможно, но чаще всего слишком сложно и дорого (см. ниже).

висимость смещения  $B$  (в ммоль/л общего холестерина) от концентрации мешающих компонентов для каждого рутинного метода:

$$\begin{aligned} (A) B &= 3,6P_{\text{пал}} - 5,2 \cdot 10^{-3}C_{\text{бел}} + 11 \cdot 10^{-3}C_{\text{бил}} - 1,9P_{\text{ар}} + 0,28; \\ (B) B &= 6,6 \cdot 10^{-3}C_{\text{бел}} - 1,6 \cdot 10P_{\text{ар}} + 0,69P_{\text{лин}} + 2,1P_{\text{ол}} - 0,11C_0 - 0,36; \\ (C) B &= 2,8 \cdot 10^{-3}C_{\text{бел}} - 7,1 \cdot 10^{-3}C_{\text{бил}} + 1,3P_{\text{лин}} + 1,2P_{\text{ол}} - 0,03C_0 - 1,11; \\ (D) B &= 2,0P_{\text{ар}} + 1,6P_{\text{лин}} - 0,37C_{\text{ол}} - 0,004C_0 - 0,73, \end{aligned}$$

где  $C_{\text{бел}}$  — концентрация общего белка, г/л;  $C_{\text{бил}}$  и  $C_{\text{тр}}$  — концентрация билирубина и триацилглицеридов соответственно, ммоль/л;  $P_{\text{ар}}$ ,  $P_{\text{лин}}$ ,  $P_{\text{ол}}$ ,  $P_{\text{пал}}$  — отношение концентрации эфиров холестерина (арахидоната, линолеата, олеата, пальмитата) к концентрации общего холестерина;  $C_0$  — концентрация общего холестерина. В уравнения включены только компоненты, влияние которых было выявлено достоверно.

Концентрация мешающих компонентов в 90% проб лежат в следующих пределах<sup>41</sup>: билирубин 1–20 мкмоль/л; триацилглицериды 0,4–2,4 ммоль/л; общий белок 64–88 г/л; общий холестерин 4,5–6,5 ммоль/л; доля арахидоната, линолеата, олеата, пальмитата холестерина соответственно равна 0,04–0,16; 0,3–0,70; 0,10–0,30; 0,06–0,14. Подставляя крайние значения этих интервалов в уравнения для смещения, получаем, что при неудачном сочетании мешающих компонентов смещение для отдельных проб может составлять для методов (A–D) (0 ÷ 14); (–6 ÷ 13); (–17 ÷ 1); (–19 ÷ 26) процентов соответственно. Изредка могут встречаться пробы и с большими смещениями. Более того, наличие в уравнениях значительных свободных членов указывает на возможное существование неизвестных нам факторов, влияющих на результаты анализа. В то же время, согласно национальным и международным стандартным документам, суммарная относительная погрешность определения общего холестерина не должна превышать 3–6%. Учет же влияния матрицы даже по известным мешающим компонентам практически невозможен. Для этого необходимо было бы в каждой пробе измерять все мешающие компоненты, а оборудования для этого в обычных лабораториях просто нет (а если и есть, то это слишком дорого).

Другой пример скрытой систематической погрешности описан в [22]. При анализе сохранившихся волос Наполеона методом нейтронно-активационного анализа были получены данные, свидетельствующие об аномально высокой концентрации в них мышьяка (что прямо указывало на возможное отравление). Позже эти результаты были подвергнуты сомнению: выяснилось, что по использованной методике фактически определялось совместное присутствие в пробе мышьяка и сурьмы. Дискуссии по этому вопросу ведутся до настоящего времени.

Скрытые систематические погрешности свойственны большинству методик, используемых для анализа объектов сложного состава. Правда, известны несколько методов, практически свободных от таких погрешностей (например, масс-спектрометрия с изотопным разбавлением), однако они чрезвычайно дороги, применимы лишь в отдельных случаях и редко используются для выполнения рутинных анализов. Такое положение сохранится и в обозримом будущем, поскольку экономические соображения заставляют все чаще применять недорогие упрощенные методики анализа (например, с применением сенсоров или тест-методы). Более того, зачастую методики с большими погрешностями вытесняют более точные!

**Пример.** Пламенная фотометрия с внутренним стандартом, высокоточная методика, широко использовавшаяся ранее при определении натрия и калия в биологических жидкостях, практически вытеснена в настоящее время методиками с использованием ион-селективных электродов, сильно уступающим ей в точности. Основная причина этого далека от науки —

<sup>41</sup> При оценке интервалов учтен разброс концентраций как внутри, так и между популяциями, то есть практически для всех реально встречающихся рутинных проб.



использование горючих газов в лечебных учреждениях запрещено из противопожарных соображений.

Как указывалось в разд. 3.1.1, наличие скрытых систематических погрешностей в большинстве случаев делает невозможным либо бессмысленным построение классической схемы передачи размера моля от первичного эталона (в качестве которого можно использовать навеску чистого вещества) к методикам анализа (которые можно рассматривать как средства измерений нижнего уровня) и стандартным образцам сложного состава. Поэтому, как правило, стандартные образцы сложного состава аттестуют путем межлабораторного эксперимента. Более того, часто стандартному образцу приписывают опорное значение (обычно это концентрация аналита) или даже несколько опорных значений, каждое из которых соответствует одной методике или группе методик. Отсюда следует, что стандартный образец выполняет функции как первичного эталона, так и образцовой меры. Все это возможно, в частности, в силу достаточно больших погрешностей методов количественного химического анализа. Подробнее об этом см. разд. 3.2 и сл.

Скрытые систематические погрешности, так же как другие виды погрешностей, исследуются в ходе *аттестации методик количественного химического анализа* (см. разд. 3.3).

Можно условно установить соответствие между метрологией «классических» средств измерений (см. гл. 1) и методик количественного химического анализа (табл. 3.1).

**Таблица 3.1.** Соответствие «классических» средств измерений и методик КХА

«Классические» средства измерений	Методики количественного химического анализа
Разработка	Разработка
Нормирование	Исследование метрологических характеристик и нормирование
Утверждение типа средства измерений	Аттестация (валидация) методики КХА
Изготовление серийного экземпляра средства измерений	Постановка методики в конкретной лаборатории
Первичная поверка	Подтверждение соответствия метрологических характеристик методики в конкретной лаборатории предъявляемым требованиям (верификация)
Периодическая поверка	Внутрилабораторный контроль качества. Внешняя оценка качества. Аккредитация и инспекционный контроль лабораторий

Каждая позиция в левом столбце имеет аналог в правом, однако методики КХА имеют много особенностей.

С учетом этого методику количественного химического анализа в целом можно рассматривать как особый объект метрологического обеспечения (с учетом того, что в классической метрологии такими объектами являются средства измерений). При этом специфика в каждом случае определяется решаемой аналитической задачей, используемыми методами, аппаратурой и т. д.

В данной книге мы делим все, что делается для получения достоверных и точных результатов химического анализа, на две части. В этой главе рассматриваются

«метрологические» аспекты (использование образцов сравнения и стандартных образцов, исследование и аттестация методик КХА и т. д.) — то есть то, без чего методики не могут применяться в принципе. В гл. 4 рассмотрены способы, используемые в лабораториях для обеспечения качества КХА.

### 3.2. Образцы сравнения и стандартные образцы

*Образцы сравнения (англ. Reference Material, RM) и стандартные образцы (англ. Certified Reference Material, CRM)*<sup>42</sup> — основа метрологии и важнейшее средство обеспечения качества количественного химического анализа [52, 55]. При этом стандартные образцы можно рассматривать как частный случай образцов сравнения.

*Образцы сравнения* — это материалы или вещества, используемые в метрологических целях, включая обеспечение качества измерений — для градуировки, аттестации методик, оценки статистических характеристик измерительного процесса, внутрилабораторного контроля и внешней оценки качества КХА и т. д.<sup>43</sup> Важнейшая отличительная особенность образцов сравнения, как и стандартных образцов, состоит в том, что содержание в них определяемого вещества (веществ) — опорное или аттестованное значение — известно, и известна погрешность этого значения (с принятой доверительной вероятностью). О способах оценки этих характеристик см. разд. 3.2.4.

Обычные пробы, ради исследования которых и проводится анализ, ниже называются *рутинными пробами*.

Образцы сравнения можно условно разделить на образцы сравнения простого и сложного состава. Под *образцами сравнения простого состава* понимаются индивидуальные чистые вещества, а также их растворы (например, чистый сульфат калия и его водный раствор). *Образцами сравнения сложного состава* являются многокомпонентные образцы, близкие по составу и свойствам к рутинным пробам сложного состава. Как правило, их готовят на основе веществ одной природы с исследуемыми объектами, подвергаемыми при этом определенной обработке, зависящей от свойств этих веществ (примеры: сухая и измельченная дерново-подзолистая почва, лиофилизированная сыворотка крови человека, титановый сплав и т. д.).

Для удобства применения образцы сравнения делят на части (расфасовывают во флаконы, ампулы, металлы и сплавы делят на отдельные бруски и т. д.). Каждую такую часть называют *экземпляром*. Образец сравнения может существовать в одном или в нескольких экземплярах, однако чаще всего из соображений удобства применения изготавливают большее их количество: максимальное число экземпляров одного образца может достигать нескольких десятков и даже сотен тысяч.

Основные требования к образцам сравнения — *стабильность* и *однородность*. Для многих образцов сравнения требуется также *адекватность* исследуемым рутинным пробам.

<sup>42</sup> Английское название Standard Reference Material, SRM — патентованное название стандартных образцов Национального института стандартов и технологий США (NIST).

<sup>43</sup> Выделяют образцы сравнения (и стандартные образцы) состава и свойств. Мы будем рассматривать только образцы сравнения (и стандартные образцы) состава веществ и материалов.



Под *стабильностью* образцов сравнения понимают следующее. Если изменение концентрации аналитов, включая форму, в которой они находятся, и наличие мешающих компонентов (если они могут влиять на результаты анализов) от момента аттестации образца сравнения до момента применения мало по сравнению с его общей погрешностью, образец считается стабильным.

Для обеспечения стабильности образцов сравнения почти всегда принимаются специальные меры. В растворы органических веществ добавляют вещества, препятствующие их окислению или имеющие бактерицидные свойства.

**Пример:** растворы глюкозы стабилизируют бензойной кислотой.

Приготавливая образцы сравнения сложного состава биологического происхождения, из них почти всегда удаляют воду (высушиванием или лиофилизацией<sup>44</sup>). Жидкости во избежание испарения герметично изолируют (например, запаивают в ампулы), а емкости с порошкообразными образцами плотно закрывают крышками для предотвращения загрязнения. Потенциально нестабильные пробы хранят и пересылают при пониженной температуре, содержат в сосудах, заполненных инертным газом, и т. д.

Стабильность образца сравнения обычно выражают в виде срока его годности, т. е. времени, в течение которого погрешность, связанная с нестабильностью, не превышает допустимого значения. Некоторые способы оценки срока годности образцов сравнения описаны ниже (разд. 3.2.2).

*Однородность (гомогенность)* образцов сравнения есть свойство, выражающееся в постоянстве значения аттестуемой характеристики для любой части образца — концентрации определяемых веществ, измеренной методами, для обеспечения качества которых предназначен образец. Характеризуют обычно *неоднородность* — степень отклонения от однородного состояния, которую рассматривают как случайную величину. Характеристикой неоднородности является *стандартное отклонение неоднородности*.

Необходимая однородность гетерофазных образцов сравнения достигается тщательным измельчением и последующим перемешиванием исходного материала. Требование однородности должно соблюдаться как в пределах каждого экземпляра образца сравнения, так и между ними. Об оценке неоднородности образцов сравнения см. разд. 3.2.3.

*Адекватность* требуется обычно от образцов сравнения сложного состава и заключается в том, что содержание определяемого вещества (веществ) в них, содержание других веществ (матрица) и физико-химические свойства близки к таким же показателям рутинных проб. Понятно, что такое определение не является строгим, поскольку в реальности анализируется множество отличающихся друг от друга рутинных проб. Подробнее об адекватности см. разд. 3.2.5.

Кроме условного деления на образцы простого и сложного состава, образцы сравнения можно классифицировать по их назначению, а стандартные образцы — и по уровню признания.

По назначению можно выделить следующие виды образцов сравнения:

<sup>44</sup> Лيوфилизация — процесс, при котором вещество быстро охлаждают до  $-10^{\circ}\text{C}$  или ниже, после чего воду удаляют под вакуумом за счет сублимации, постепенно повышая температуру.

- а) образцы для оценки точности анализа;
- б) образцы для градуировки;
- в) образцы-добавки;
- г) суррогаты.

*Образцы для оценки точности анализа* составляют большинство образцов сравнения. Как правило, их анализируют так же, как рутинные пробы, и по результатам их исследования судят о правильности, а иногда и о прецизионности анализа в целом.

*Образцы для градуировки*, как следует из названия, используют при градуировке. В терминах количественного химического анализа это зависимость аналитического сигнала (например, показаний используемого средства измерений) от концентрации определяемого вещества (веществ), см. разд. 3.3.1.

*Образцы-добавки* предназначены для контроля правильности результатов анализа или для градуировки с использованием метода «введено-найдено» либо для нахождения результата методом добавок (см. разд. 3.3.4). Основное требование, предъявляемое к таким образцам, состоит в возможности их равномерного распределения в пробах, аналогичных рутинным, причем химическая форма и фазовое состояние («форма вхождения») вводимых аналитов не должны отличаться от таковых для аналитов, изначально присутствующих в пробах.

*Суррогат* в химическом анализе представляет собой вещество, отличающееся от определяемого, но которое в процессе анализа ведет себя так же, или практически так же, как определяемый компонент. Классические суррогаты — изотопзамещенные соединения, заведомо ведущие себя так же, как определяемые вещества в силу близости их химических и физических свойств к аналитам.

**Пример.** При определении концентрации холестерина в сыворотке крови человека методом масс-спектрометрии с изотопным разбавлением используются изотопзамещенные суррогаты.

Пробу сыворотки крови обрабатывают щелочью для гидролиза эфиров холестерина, затем в нее вводят (обычно в форме спиртового раствора) навеску холестерина, в котором часть атомов замещена изотопами; чаще всего часть атомов водорода заменяют на дейтерий — это и есть суррогат. Далее экстрагируют холестерин неполярным растворителем (обычно гексаном) и вводят пробу в хромато-масс-спектрометр. Искомую концентрацию холестерина находят по соотношению молекулярных пиков с массами, соответствующими природному и изотопзамещенному холестерину. В этом случае (как это чаще всего и бывает) суррогат используется для градуировки. Данный вариант использования суррогата называю также «методом внутреннего стандарта».

Использовать другие (не изотопзамещенные) виды суррогатов необходимо с большой осторожностью и только после тщательных исследований, подтверждающих несущественность различий между поведением суррогата и аналита в ходе анализа.

Образцы сравнения, утвержденные (рекомендованные к использованию) официальными организациями, называют *стандартными образцами состава*, или просто *стандартными образцами (СО)*<sup>45,46</sup>. К стандартным образцам, как к частному случаю образцов сравнения, относится все, что сказано выше об образцах сравнения.

<sup>45</sup> Стандартные образцы сложного состава называют также «стандартными образцами состава сложных матриц» [56].

<sup>46</sup> Иногда образцы сравнения называют «стандартными образцами», а официально признанные образцы сравнения — аттестованными или сертифицированными стандартными образцами». Это создает путаницу.



По уровню признания можно выделить (в порядке понижения этого уровня) следующие категории образцов сравнения.

*Международные или национальные стандартные образцы* вместе с документацией утверждаются (рекомендуются к использованию) международными или национальными организациями. При этом международные СО могут разрабатываться как этими организациями, так и другими институтами — важен лишь уровень их признания. Естественно, авторитетность международного или национального стандартного образца определяется авторитетом утвердившей его организации. В России узаконены [52] следующие виды СО этого уровня признания:

- стандартные образцы утвержденных Федеральным агентством по техническому регулированию и метрологии типов (ГСО);
- межгосударственные стандартные образцы (МСО), признанные Межгосударственным советом по стандартизации, метрологии и сертификации (МГС);
- стандартные образцы, признанные Комитетом Евро-Азиатского сотрудничества государственных метрологических учреждений (СО КООМЕТ).

Подробнее вопрос об утверждении стандартных образцов рассмотрен в разд. 3.2.6.

*Стандартные образцы, утверждаемые на уровне организаций, корпораций, объединений, ведомств и других юридических лиц.* Они разрабатываются и производятся этими юридическими лицами, которые делают это самостоятельно и сами проводят аттестацию этих образцов. Отдельные аббревиатуры имеют следующие виды СО [52]:

- СО государственных научных метрологических институтов (СОГНМИ);
- отраслевые стандартные образцы (ОСО);
- стандартные образцы предприятий (СОП).

СО этого вида применяются вне сферы государственного регулирования обеспечения единства измерений.

Помимо образцов сравнения для оценки ряда характеристик аналитических методик, при которой точное знание концентрации определяемых веществ не обязательно, а также для внутрилабораторного контроля, из экономии часто применяются однородные и стабильные (в течение времени их использования) образцы, адекватные рутинным пробам. Такие образцы, отличающиеся от образцов сравнения сложного состава тем, что концентрация определяемых веществ в них известна лишь ориентировочно, будем называть *контрольными материалами* или *контрольными образцами*.

В практике создания стандартных образцов можно выделить стадии их исследования (*валидация*) и утверждения. Часто их рассматривают вместе, называя это *аттестацией стандартных образцов*.

Разработке подлежит *тип стандартного образца*, т. е. совокупность стандартных образцов, изготовляемых из одних и тех же веществ по одной и той же технологии, содержание определяемых веществ в которых устанавливается одним и тем же способом.

**Пример.** Разработан и утвержден тип стандартного образца (ГСО 9968-2011. СО состава смеси молочной), в котором аттестуется содержание тяжелых металлов — меди, кадмия и мышьяка, предназначенный для контроля точности определения этих металлов в пищевых

продуктах. Это значит, что для этого СО утверждены требования, технология его приготовления, способы контроля качества и установления концентрации аналитов, и в соответствии с ними в дальнейшем можно изготовить сколько угодно стандартных образцов этого типа.

На стадии разработки определяются назначение и требования к будущему стандартному образцу: состав, уровень признания (категория), требования к однородности, стабильности, погрешности установления содержания определяемых веществ и т. д. Далее разрабатывается и реализуется технология приготовления стандартного образца, позволяющая достичь требуемых характеристик. Разработка стандартных образцов, особенно стандартных образцов сложного состава, является областью высоких технологий. Как правило, стоимость разработки стандартного образца очень высока. Это обуславливает высокие цены на них. Во многих случаях разработка новых типов стандартных образцов финансируется государственными организациями в форме бюджетных субсидий, грантов и т. д.

Исследование стандартного образца может проводиться как в ходе разработки типа стандартного образца, так и при изготовлении нового стандартного образца ранее разработанного типа. Обычно исследование включает оценку стабильности и неоднородности стандартного образца, а также оценку содержаний определяемых веществ и их погрешностей (неопределенностей). Соответственно, результатом исследования является установление содержания определяемого вещества (веществ) в стандартном образце и его погрешности (неопределенности), срока хранения и других метрологических характеристик.

Утверждение (признание) стандартного образца есть юридическая процедура, в ходе которой ему придают официальный статус. В странах СНГ основным нормативным документом, регламентирующим разработку, выпуск и применение стандартных образцов, является ГОСТ 8.315 [56], а утверждение типа стандартного образца проводится в соответствии с законом [2].

Все, что говорится ниже о стандартных образцах, относится также и ко всем образцам сравнения (за исключением вопросов, связанных с их утверждением).

Часто используются не аттестованные, то есть не имеющие опорного значения материалы. Они наряду с образцами сравнения и стандартными образцами применяются при организации межлабораторных экспериментов (см. разд. 3.6), для внутрилабораторного контроля (см. разд. 4.5) и т. д. Для того чтобы охватить этот вид материалов, мы будем использовать общий термин «*контрольные материалы*», подразумевая любые однородные, стабильные и адекватные материалы, как аттестованные, так и не аттестованные.

### 3.2.1. Погрешность стандартного образца

Важнейшими величинами, характеризующими стандартный образец состава вещества, являются концентрации определяемых веществ (аналитов) и общая погрешность этих концентраций.

Далее для простоты мы будем рассматривать процедуру исследования СО так, как будто в данном стандартном образце определяется только один аналит. Конечно, все сказанное справедливо для каждого аналита в случае, когда в СО определяется концентрация нескольких веществ.



Базовым требованием, предъявляемым к любому стандартному образцу, является требование к общей допустимой (нормируемой) погрешности  $\Delta_{\text{CO}}$ , с которой известна концентрация (концентрации) определяемого в нем вещества (веществ). Естественно, эта погрешность может быть различной для разных определяемых веществ. Согласно принципу пренебрежения малыми погрешностями (см. разд. 2.6.1), общая допустимая погрешность стандартного образца  $\Delta_{\text{CO}}$  должна быть мала по сравнению с погрешностями аналитической методики, для контроля точности которой предназначен стандартный образец:

$$\Delta_{\text{CO}} < 1/3 \Delta_{\text{д}}, \quad (3.1)$$

где  $\Delta_{\text{д}}$  — общая нормируемая погрешность методики химического анализа. *Здесь и ниже предполагается, что допустимая погрешность анализа  $\Delta_{\text{д}}$ , которая может определяться различными способами (из метрологических характеристик самой методики, включая используемую аппаратуру; исходя из требований к результатам анализа и т. д.) задана. Мы не обсуждаем здесь возможные составляющие этой погрешности.*

Это жесткое и не всегда выполнимое требование. Действительно, при относительной суммарной погрешности методики ( $\pm 6\%$ ) относительная погрешность стандартного образца не должна превышать ( $\pm 2\%$ ), и это включая все ее компоненты: погрешность определения концентрации, погрешности, связанные с неоднородностью, нестабильностью и т. д.

Обычно принимают, что различные источники погрешности независимы. Тогда, согласно закону сложения погрешностей, можно записать

$$\Delta_{\text{CO}}^2 = \xi_a^2 + \xi_h^2 + \xi_s^2 + \xi_{\text{MA}}^2, \quad (3.2)$$

где  $\Delta_{\text{CO}}^2$  — дисперсия, характеризующая общую погрешность стандартного образца,  $\xi_a^2$ ,  $\xi_h^2$ ,  $\xi_s^2$  и  $\xi_{\text{MA}}^2$  — соответственно, дисперсии, характеризующие погрешности, определяемые: процедурой установления средней концентрации аналита, неоднородностью, нестабильностью и неадекватностью СО.

Максимально допустимые частные погрешности  $\xi_a^2$ ,  $\xi_h^2$ ,  $\xi_s^2$  и  $\xi_{\text{MA}}^2$  выбирают так, чтобы их сумма, т. е. общая погрешность, не превышала заданную величину. Однако если какая-либо частная погрешность не может быть уменьшена до требуемого значения, допускают увеличение общей погрешности. Другими словами, в случае невозможности выполнения требования (3.1) добавляются минимально возможной по техническим и экономическим соображениям погрешности. Такая же ситуация возникает, когда стандартный образец не предназначен для конкретной цели, например, для контроля точности определенной методики (методик) КХА. В таких случаях стараются просто добиться минимально возможной погрешности СО. Подробнее этот вопрос рассмотрен в [55]. Некоторые способы оценки составляющих погрешности приведены ниже.

Как и при любом сложении погрешностей, для оценок дисперсий формула (3.2) становится приближенной.

### 3.2.2. Оценка срока годности стандартных образцов

Срок годности стандартного образца обычно оценивают экспериментально. Если в стандартном образце определяют несколько показателей (аналитов), часто из экономии оценивают срок годности по одному или нескольким показателям, которые полагают наименее стабильными.

Для оценки срока годности необходимо найти зависимость концентрации определяемого вещества в образце от времени (при условии, что стандартный образец хранится при паспортных условиях) и установить, когда изменения концентрации достигнут максимально допустимой величины. Эту величину обычно принимают равной  ${}^2/_{3}\Delta_{\text{CO}}$ .

*Простейший способ* оценки срока хранения заключается в периодических измерениях концентрации определяемого вещества (веществ) в стандартном образце, хранящемся в стандартных для него условиях. Последовательность действий такова.

1. Выбирают продолжительность (общее время)  $\Theta$  и периодичность исследования. Время, в течение которого проводится исследование, не должно быть меньше половины предполагаемого срока годности данного образца. Измерения концентрации обычно проводят через равные промежутки времени

$$0; \frac{\Theta}{N-1}; \frac{2\Theta}{N-1}; \dots; \frac{n\Theta}{N-1}; \dots; \Theta, \quad (3.3)$$

где  $n$  ( $0, 1, \dots, N-1$ ) — номер момента, в который проводится измерение,  $N$  — общее число моментов, в которые проводятся измерения. При такой записи при расчетах для обозначения времени можно использовать  $n$ .

Периодичность, с которой проводятся измерения, может быть различной, однако слишком частые измерения бессмысленны. Другими словами, величина  $N$  не должна быть большой.

2. Выбирают число измерений  $K$ , выполняемых в каждый момент времени  $n$  (несколько измерений проводят для повышения точности). Вообще говоря, количество измерений должно выбираться таким образом, чтобы оценки коэффициента  $\beta$  (см. ниже, модель 3.5 и сл.) имели бы погрешность, позволяющую достоверно зарегистрировать изменения концентрации в течение предполагаемого срока хранения, не превышающие  ${}^2/_{3}\Delta_{\text{CO}}$ . На практике число измерений  $K$  выбирают исходя из оценки стандартного отклонения используемого метода анализа  $s_{\text{ан}}$

$$K \geq (s_{\text{ан}}/2\Delta_{\text{CO}})^2. \quad (3.4)$$

3. Проводят эксперимент, измеряя концентрацию определяемого вещества в образце в выбранные моменты времени. Обозначим получаемые результаты как  $C_{n,k}$ , где  $k = 0, 1, \dots, K$  — номер измерения, выполненный в момент времени  $n$ . Результаты обрабатывают в рамках модели регрессионного анализа

$$C_{n,k} = \alpha + \beta n + \xi, \quad (3.5)$$

где  $\alpha = C_{n=0}$  — концентрация в начальный момент времени  $n = 0$ ;  $\beta$  — коэффициент, характеризующий скорость изменения концентрации;  $\xi \in N(0, \sigma_{\xi}^2)$  случайная по-



грешность измерений. Линейную модель (3.5) выбирают потому, что она хорошо описывает зависимость концентрации от времени при небольших ее изменениях.

Модель (3.5) можно записать также в виде

$$C_{n,k} - C_{n=0} = \beta n + \xi. \quad (3.6)$$

4. Проверяют гипотезу  $H_0: (\beta = 0)$ . Если она не отвергается, стандартный образец считают стабильным в течение срока выполнения эксперимента и устанавливают срок годности, равный удвоенной продолжительности эксперимента. Если гипотеза  $H_0$  отвергается с принятым уровнем доверительной вероятности, находят оценку  $b$  коэффициента  $\beta$  и рассчитывают срок годности стандартного образца как частное от деления максимально допустимой погрешности нестабильности, составляющей  $^{2/3}\Delta_{CO}$ , на  $b$ . Методы расчетов приведены в разд. 2.7.2. Срок годности в любом случае ограничивается (обычно максимум 5 годами).

Такой способ имеет два основных недостатка.

1. Методика, которой проводятся измерения, может быть подвержена медленному дрейфу, который в рамках модели (3.5) не отличим от изменений стандартного образца. Выходом было бы проведение измерений в нескольких лабораториях (межлабораторный эксперимент), однако такая процедура слишком дорога и практически не используется. Обычно для оценки стабильности стараются применять наиболее устойчивые методики, для которых риск медленного дрейфа минимален, хотя и это не всегда помогает.

2. Продолжительность проверки стабильности стандартного образца очень велика — до двух лет, и мало кто соглашается так долго ждать результата.

Для повышения надежности применяют *метод сравнительных измерений*. Его используют, когда с помощью специальных приемов можно обеспечить условия хранения, заведомо гарантирующие стабильность образца (например, путем глубокого замораживания, помещения образца в инертную атмосферу и т. д.). В этом случае часть стандартного образца хранят в условиях, гарантирующих его стабильность, и для этой части проводят те же измерения, что и для остальной части стандартного образца (хранящегося в обычных условиях). При этом считается оптимальным в каждый момент времени  $n$  чередовать измерения, относящиеся к образцам, хранившимся при разных условиях. При таком способе измерений расчеты проводят в рамках модели

$$C_{n,k} - C_{n,k}(\text{const}) = \beta n + \xi. \quad (3.7)$$

где  $C_{n,k}$  — результаты измерения для стандартного образца, хранящегося в стандартных условиях;  $C_{n,k}(\text{const})$  — результаты измерений для части стандартного образца, хранящегося в условиях, гарантирующих его стабильность,  $\xi$  — случайная погрешность. В методе сравнительных измерений практически устраняется возможность артефактов, связанных с изменениями используемого метода анализа во времени.

*Метод ускоренного старения* [57] используется для сокращения времени оценки срока годности. Он основан на том, что для большинства стандартных образцов нестабильность, определяемая в основном протеканием химических реакций, сильно зависит от температуры. Метод ускоренного старения заключается в исследовании частей стандартного образца, хранящихся при повышенных температурах.

Рассмотрим применение метода ускоренного старения для наиболее частого случая, когда концентрация аналита уменьшается во времени за счет протекания химической реакции первого порядка (или других явлений, при которых изменение концентрации описывается такой же кинетической кривой). Такой процесс часто называют деградацией, и его при температуре  $T$  можно охарактеризовать константой скорости деградации  $kT$ . Предполагается, что во всем интервале температур, в котором проводится исследование, зависимость константы скорости деградации от температуры описывается уравнением Аррениуса

$$kT = A \exp(-E_a/RT), \quad (3.8)$$

где  $A$  — предэкспоненциальный множитель;  $E_a$  — энергия активации;  $R$  — универсальная газовая постоянная. Фактически такая гипотеза эквивалентна предположению о том, что химический механизм деградации один и тот же во всем исследуемом интервале температур. Этот интервал ограничен сверху. Например, для биологических проб редко используют температуру выше 40–50 °С, поскольку при более высоких температурах денатурируются белки, в том числе ферменты, и механизм деградации изменяется.

Измерения по методу ускоренного старения проводят следующим образом. Отделяют от стандартного образца несколько, желательно не меньше 3–4, частей (обычно каждая часть состоит из нескольких экземпляров) и хранят каждую из них при определенной температуре  $T$ . При этом температурный интервал стараются задать по возможности широким (с учетом сохранения механизма реакций). Через равные промежутки времени каждую часть анализируют; число измерений для каждой части выбирается на основании соотношения (3.4). Обычно все измерения, выполненные в одно время, проводят в одной аналитической серии (см. разд. 3.3.1).

Обозначим результат  $k$ -го измерения в момент времени с номером  $n$  для части стандартного образца, хранящегося при температуре  $T$ , как  $C_{n,k}(T)$ . Полагая, как и раньше, зависимость концентрации от времени при одной и той же температуре линейной, обрабатываем результаты для части СО, хранящейся при температуре  $T$ , в рамках модели регрессионного анализа

$$C_{n,k}(T) = \alpha + \beta n + \xi, \quad (3.9)$$

где  $\alpha$  — отрезок, отсекаемый на оси ординат,  $\beta$  — тангенс угла наклона прямой (представляет собой константу скорости деградации при температуре  $T$ ),  $\xi$  — нормально распределенная случайная погрешность. Соответственно, оценка величины  $\beta$  есть оценка  $kT$ .

Далее находят зависимость  $kT$  от  $1/T$ . При выполнении уравнения Аррениуса (3.8) эта зависимость линейна; отклонения от линейности указывают на изменение механизма деградации стандартного образца. По полученному графику и определяют константу скорости реакции деградации  $k$  при температуре, при которой предполагается хранить СО. Зная  $k$ , можно рассчитать концентрацию аналита  $C_t$  в момент времени  $t$  по формуле

$$\lg C_t = \lg C_0 - (kt/2,303), \quad (3.10)$$

где  $C_0$  — концентрация в начальный момент времени.



Учитывая, что максимально допустимое изменение концентрации при хранении не должно превышать  $2/3 \Delta_{CO}$ , срок хранения  $t_{\max}$  не должен превышать

$$t_{\max} \leq \frac{2,303}{k} \lg \left( \frac{2}{3} \frac{\Delta_{CO}}{C_0} \right). \quad (3.11)$$

В упрощенном варианте метода ускоренного хранения исследуют пробы, хранившиеся при двух температурах — повышенной и стандартной для данного стандартного образца. В этом случае при расчетах используют правило Вант-Гоффа: при повышении температуры на  $10^\circ\text{C}$  скорость реакции возрастает в 2–3 раза. Это менее трудоемко, но и менее точно.

Возможна также комбинация метода ускоренного хранения и метода сравнительных измерений.

Часто результаты эксперимента по оценке срока хранения подвергают процедуре экспоненциального сглаживания [58], близкой к применяемой при построении контрольных карт геометрических движущихся средних (см. разд. 4.5.1.3), что позволяет несколько снизить случайную погрешность за счет возможного увеличения систематических погрешностей.

Может оказаться, что срок хранения, оцененный приведенными способами, слишком мал. В таких ситуациях, если удастся, устанавливают зависимость концентрации от времени (при хранении образца в определенных условиях) и вносят ее в паспорт стандартного образца. Так, например, указывается концентрация глюкозы в некоторых стандартных образцах лиофилизированной сыворотки крови. Другой способ — допускают увеличение погрешности стандартного образца и учитывают погрешность нестабильности при ее расчете по формуле (3.2).

### 3.2.3. Оценка неоднородности стандартных образцов

Неоднородность химического состава, т. е. неравномерность распределения индивидуальных веществ (компонентов), — естественное свойство большинства веществ и материалов. Конечно, при приготовлении стандартных образцов стараются уменьшить неоднородность — главным образом путем грубого и тонкого измельчения, рассеивания, перемешивания — однако полной однородности добиться не удается. Исключений немного: это истинные растворы, ряд биологических жидкостей и др. *Фактически все твердые стандартные образцы неоднородны.*

Неоднородность вещества стандартного образца приводит к различным результатам анализа проб, взятых из разных его частей, т. е. к возникновению погрешности за счет неоднородности. Заметим, что химиков-аналитиков интересует в основном связанная с неоднородностью погрешность анализа, а не ее причины (неоднородность распределения определяемых веществ, гранулометрического состава, физико-химических свойств и т. д.). Наиболее часто, когда стандартный образец делится на экземпляры (фасуется), необходимо учитывать возможный разброс концентрации между экземплярами — «межэкземплярное» или «межфасовочное» стандартное отклонение [59]. Как и при оценке срока хранения стандартных образцов, при оценке неоднородности часто измеряют концентрацию не

всех определяемых веществ, а лишь тех из них, распределение которых в массе СО *a priori* наиболее неоднородно.

Для оценки неоднородности СО проводят специальный эксперимент. При этом учитывают назначение образца и с этой точки зрения выделяют два основных случая: 1) стандартный образец предназначен для одной определенной цели, например для оценки качества конкретных методик анализа, и (минимальная) масса его, которая будет отбираться для анализа, известна; 2) стандартный образец может использоваться для различных целей, и объем пробы при этом может быть различен (и заранее неизвестен), а задается максимально допустимая погрешность неоднородности СО.

**В первом случае** проводится эксперимент, цель которого — проверить гипотезу о том, что погрешность неоднородности СО  $\xi_h$  при заданной массе навески мала по сравнению с общей погрешностью стандартного образца  $\Delta_{\text{СО}}$ . Если эта гипотеза отвергается, то оценивают  $\xi_h$  для учета ее вклада в общую погрешность согласно (3.2). Эксперимент планируется и выполняется следующим образом.

1. Оценивают стандартное отклонение сходимости методики анализа  $s_r$  с использованием однородного вещества. Это наиболее сложная задача. Дело в том, что при любом плане эксперимента по оценке неоднородности СО разброс результатов определяется суммой случайной погрешности анализа и погрешности за счет неоднородности СО. Поэтому погрешность анализа (поскольку эксперимент обычно проводится в одних условиях, то он характеризуется стандартным отклонением повторяемости (сходимости), см. разд. 3.3.1) необходимо оценивать отдельно, и делать это надо весьма точно. Идеальной является ситуация, когда  $s_r$  оценивают непосредственно в ходе эксперимента, например, измеряя совместно с пробами исследуемого СО заведомо однородное вещество, близкое к нему по свойствам. К сожалению, такого вещества зачастую просто не существует. Поэтому применяют паллиативные способы — оценивают разброс между результатами параллельных определений после растворения пробы (анализ нескольких аликвот), анализируют большие навески и т.д. Во всех подобных случаях анализ происходит несколько по-другому, чем при оценке неоднородности СО, и вопрос о степени близости оцененной величины  $s_r$  к сходимости методики анализа при эксперименте по оценке неоднородности СО должен рассматриваться отдельно в каждом случае.

2. Устанавливают число проб (навесок заданной массы), которые необходимо отобрать для исследования. При этом исходят из двух положений:

- во-первых, общая дисперсия результатов исследования неоднородности СО  $s_{\text{общ}}^2$  равна сумме дисперсии повторяемости (сходимости) анализа и общей погрешности за счет неоднородности

$$s_{\text{общ}}^2 \cong s_r^2 + s_h^2, \quad (3.12)$$

где  $s_r^2$  — дисперсия повторяемости (сходимости) метода,  $s_h^2$  — дисперсия неоднородности СО (оценка  $\xi_h^2$ );

- во-вторых, необходимо достоверно выявить  $s_h = 0,3\Delta_{\text{СО}}$ . Это означает, что при проверке гипотезы  $H_0: (s_{\text{общ}}^2 = s_r^2)$  она должна отвергаться при  $s_h = 0,3\Delta_{\text{СО}}$ . В этом случае можно рассчитать  $s_h^2$  как

$$s_h^2 = (s_{\text{общ}}^2 - s_r^2). \quad (3.13)$$



Гипотеза  $H_0: (s_{\text{общ}}^2 = s_r^2)$  отвергается при  $F = s_{\text{общ}}^2 / s_r^2 > F(\bar{P}, f_{\text{общ}}, f_r)$ , где  $\bar{P}$  – выбранный уровень доверительной вероятности (обычно 0,95),  $f_{\text{общ}}$  – число степеней свободы для результатов оценки  $s_{\text{общ}}^2$ ,  $f_r$  – число степеней свободы при оценке неоднородности (в обоих случаях это число измерений минус 1). Очевидно, число измерений  $N$  должно выбираться таким, чтобы

$$F(\bar{P}, f_{\text{общ}}, f_r) \cong s_{\text{общ}}^2 / (s_{\text{общ}}^2 - 0,1\Delta_{\text{CO}}^2). \quad (3.14)$$

Поскольку в большинстве случаев  $s_{\text{общ}}^2$  и  $\Delta_{\text{CO}}^2$  – величины одного порядка, трудно убедиться, что для выполнения условия (3.14) необходимо выполнить от нескольких десятков до нескольких сотен измерений.

3. Отбирают  $N$  навесок заданной массы и анализируют каждую из них, получая выборку объемом  $N$ . При этом возможны два случая.

*а.* Исходя из способа приготовления стандартного образца, можно предположить, что в нем отсутствуют крупные (относительно отбираемой пробы) зоны с существенно различающейся концентрацией аналита. Тогда полагают, что концентрация варьирует более или менее случайным образом по всему объему стандартного образца. В этом случае навески можно отбирать случайным образом или, что практически то же самое, равномерно. Последнее означает, что все вещество стандартного образца условно разбивается на  $N$  одинаковых зон и из каждой зоны отбирается одна проба. Далее пробы анализируют и проводят расчеты по формулам (3.12, 3.13).

*б.* Другой подход применяют, когда предполагают, что концентрация определяемого вещества в стандартном образце имеет два вида неоднородностей: одна – местная, аналогична рассмотренной выше, другая – *зональная*, когда в стандартном образце можно выделить большие по сравнению с массой отбираемой пробы зоны, существенно различающиеся по средней концентрации. Такими зонами могут быть различные части монолитных стандартных образцов (например, сплавов, в брусках которых близкие к центру области могут отличаться от периферийных), фасовки СО порошкообразных веществ (например, горных пород, фракции которых из-за сегрегации содержатся в разных фасовках в разных соотношениях), экземпляры лиофилизированных биологических проб (из-за разной остаточной влажности и других причин) и т. д. Зональная неоднородность СО должна учитываться при планировании эксперимента. Рассмотрим, как это делается для случая существенно отличающихся между собой экземпляров.

Вначале отбирают случайным образом некоторое число экземпляров (фасовок). Далее из каждого экземпляра фасовки отбирают несколько проб<sup>47</sup> и в каждой из них измеряют концентрацию определяемого вещества (веществ). Результаты эксперимента могут быть описаны в рамках модели дисперсионного анализа, аналогичной модели (2.52)

$$X_{ji} = \alpha + \alpha_j + \varepsilon_{ji}. \quad (3.15)$$

<sup>47</sup> Мы не приводим способы определения объема выборок, которые рассматриваются в теории планирования эксперимента (см. также разд. 2.6.4). Заметим только, что в большинстве случаев невыгодно брать больше 2-3 проб из каждой фасовки.

Здесь  $X_{ji}$  – результат  $i$ -го измерения в  $j$ -й фасовке (экземпляре) СО,  $\alpha$  – средняя концентрация в стандартном образце,  $\alpha_j$  – эффект фасовки;  $\varepsilon_{ji}$  – случайная погрешность, соответствующая сумме случайной погрешности анализа и вариации концентрации аналита от пробы к пробе в пределах фасовки. Как обычно, предполагают эффекты  $\alpha_j$  и  $\varepsilon_{ji}$  имеющими нормальное распределение с нулевыми средними значениями и дисперсиями  $\sigma_\alpha^2$  и  $\sigma_\varepsilon^2$  соответственно. Методами, описанными в разд. 2.6.2, проверяют гипотезу об отсутствии «межэкземплярного» разброса  $H_0: (\alpha_j = 0)$  и, если она отвергается, находят оценки  $\sigma_\alpha^2$  и  $\sigma_\varepsilon^2$  ( $s_\alpha^2$  и  $s_\varepsilon^2$  соответственно). Далее рассчитывают дисперсию погрешности неоднородности внутри экземпляров  $s_{в/э}^2 = (s_\varepsilon^2 - s_r^2)$ , соответствующую разбросу концентраций в пределах одного экземпляра, и общую дисперсию погрешности неоднородности по формуле

$$s_h^2 = (s_{в/э}^2 + s_\alpha^2). \quad (3.16)$$

Если гипотеза  $H_0: (\alpha_j = 0)$  не отвергается, все результаты измерений  $X_{ji}$  рассматривают как простую выборку, и проводят расчеты по формуле (3.13), используя в качестве величины  $s_{общ}^2$  дисперсию всех  $X_{ji}$ .

Заметим, что при выполнении анализа имеет смысл расставлять пробы в случайном порядке или хотя бы чередовать пробы из разных экземпляров (фасовок), иначе результаты оценки неоднородности СО могут искажаться за счет дрейфов методики анализа [59].

Во втором случае, когда стандартный образец может использоваться для различных целей и масса пробы заранее неизвестна, обычно находят зависимость погрешности неоднородности СО от массы навески. Для этого оценивают погрешность неоднородности вышеописанными способами при нескольких значениях массы отбираемой пробы. Обычно относительное стандартное отклонение неоднородности стандартного образца  $RSD_h = s_h / \bar{X}$  связано с массой пробы  $m$  соотношением

$$RSD_h \cong \frac{K}{\sqrt{m}}, \quad (3.17)$$

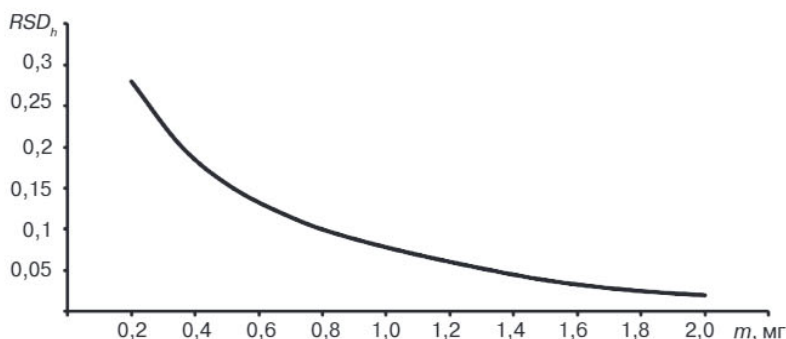


Рис. 3.2. Типичная зависимость относительного стандартного отклонения неоднородности  $RSD_h$  от массы навески  $m$

где  $K = \text{const}$ . Строят зависимость  $s_h$  от  $m$  (рис. 3.2) или непосредственно оценивают коэффициент  $K$ . После этого легко оценить  $s_h$  при любой массе отбираемой пробы, а также найти минимальную навеску, при которой погрешность неоднородности будет соответствовать условию  $s_h < 0,3\Delta_{\text{CO}}$ .

О способах оценки неоднородности стандартных образцов см. также [60].

### 3.2.4. Способы установления содержания аналитов в стандартных образцах

По определению в образцах сравнения и стандартных образцах должно быть известно содержание определяемых веществ (аналитов), и его необходимо устанавливать для каждого конкретного СО. Можно выделить четыре основных способа установления содержания определяемых веществ в стандартных образцах. Эту величину называют по-разному. В документах по аттестации СО это — «*аттестованное значение*», [52], а при использовании стандартных образцов — «*принятое опорное значение*» или просто «*опорное значение*» [27]). Заметим, что в большинстве случаев общая погрешность стандартного образца, характеризующая дисперсией  $\Delta_{\text{CO}}^2$ , определяется в основном именно погрешностью нахождения аттестованного значения

1. Оценка по *процедуре приготовления* используется в тех случаях, когда стандартный образец готовится путем смешивания (растворения, сплавления) двух или нескольких компонентов и когда нет оснований опасаться неадекватности образцов. Другими словами, это СО простого состава.

**Пример** — стандартные образцы водных растворов солей металлов, широко применяемые для градуировки и контроля точности методик определения концентрации этих металлов.

При оценке содержания по процедуре приготовления состав стандартного образца рассчитывают исходя из концентрации веществ в смешиваемых компонентах и пропорций, в которых они смешиваются. Погрешность концентрации каждого аналита вычисляют как сумму погрешностей, с которыми известно его содержание в компонентах смеси, и погрешностей, возникающих на каждом этапе приготовления (например, погрешность дозирования компонентов), по принципам сложения погрешностей (см. разд. 2.6.1).

Другие способы связаны с анализом стандартного образца.

2. *Анализ в одной лаборатории*, как следует из названия, состоит в том, что концентрация аналита определяется в стандартном образце (обычно многократно) в одной лаборатории. Значение концентрации аналита в СО рассчитывается как среднее результатов этих измерений (после исключения грубых промахов). Погрешность нахождения аттестованного значения при этом есть погрешность измерений. Такой способ, обычный в «классической» метрологии, в практике количественного химического анализа применяется в исключительных случаях. Это связано со сложностью процедур химического анализа и, как следствие, с наличием межлабораторных погрешностей — одни и те же методики при реализации их в разных лабораториях дают несколько отличающиеся между собой результаты. Еще более существенно, что каждая методика может иметь собственные скрытые систематические погрешности (см. выше).



Исключение составляют уникальные методики, имеющие малые систематические (и случайные) погрешности. В качестве иллюстрации можно привести использование масс-спектрометрии с изотопным разбавлением для оценки содержания кальция, общего холестерина, глюкозы и ряда других веществ в стандартных образцах биологических жидкостей (сыворотка крови, моча).

**Пример.** Суть метода масс-спектрометрии с изотопным разбавлением состоит в том, что в анализируемую пробу вводится точно известное количество практически отсутствующего в природе изотопа определяемого вещества (для сложных соединений — изотопзамещенного производного). Далее проводятся все необходимые для выделения определяемого вещества операции (осаждение мешающих компонентов, экстракция, хроматографическое разделение и т. д.), после чего фракция, содержащая определяемое вещество, вводится в масс-спектрометр. Концентрация рассчитывается по отношению молекулярных пиков природных и введенных изотопов (изотопзамещенных соединений). Поскольку химические и физико-химические свойства изотопов практически идентичны, основные погрешности промежуточных стадий анализа отсутствуют. Подобные методики имеют общую относительную погрешность не более 1%. Для надежности правильность применяемой методики все равно подтверждают в 2–3 лабораториях.

**3. Межлабораторный эксперимент.** Этот наиболее часто используемый способ установления содержания определяемых веществ в стандартном образце заключается в том, что части стандартного образца (обычно состоящие из нескольких экземпляров) передаются для анализа в как можно большее (но не менее чем в 8) число лабораторий. Результаты после отброса грубых промахов усредняются, и среднее значение принимается в качестве аттестованного (опорного). Подробно о межлабораторном эксперименте см. разд. 3.6.

При этом неявно принимают 1) что используемый метод (методы) как таковые не имеют существенных систематических погрешностей и 2) что результаты, полученные в разных лабораториях, независимы. Другими словами, предполагается, что погрешности измерений в разных лабораториях разнонаправлены: если погрешность для одних лабораторий положительна, то для других — отрицательна. Тогда погрешности, являющиеся систематическими для каждой лаборатории, при межлабораторном эксперименте оказываются случайными и усредняются. Грубо говоря, предполагается, что все не могут ошибаться «в одну сторону», а метод сам по себе не привносит общую систематическую погрешность в результаты всех лабораторий. Последнее может не выполняться, например, потому, что существенная часть лабораторий имеет общие источники реактивов и материалов для градуировки, работает по одной и той же модификации методики, и по другим причинам. Поэтому при организации межлабораторного эксперимента стараются не только увеличить насколько возможно число лабораторий, использующих разные методы, но и привлечь совершенно не связанные между собой лаборатории (из разных регионов или даже разных стран, разных отраслей промышленности и т. д.).

Если используемый лабораториями метод имеет систематическую погрешность при анализе исследуемого образца, узнать об этом только из результатов данного эксперимента невозможно. Однако если лаборатории используют разные методы (методики), имеющие разные систематические погрешности, это может быть видно из гистограммы результатов эксперимента, которая может иметь не один, а два или даже несколько максимумов (би- или полимодальное распределение). В таком



случае простое усреднение результатов некорректно. Конечно, причиной данного эффекта может быть и неоднородность исследуемого контрольного материала, однако это очень редкий случай. Чаше причиной все же является применение лабораториями разных методов (методик) анализа, источников образцов сравнения для градуировки и т. д. Дадим несколько типичных примеров.

**Примеры.** При определении тяжелых металлов в пищевых продуктах отличаются (в среднем) результаты, полученные с применением атомно-абсорбционного и атомно-эмиссионного методов анализа. По многим химическим элементам различаются результаты, полученные при сожжении проб в открытом и закрытом (бомбы) объеме.

При определении органических веществ в стандартных образцах сыворотки крови различающиеся результаты дают химические и ферментные методы, методики с градуировкой по образцу сравнения сложного состава и по растворам определяемых веществ в воде.

При анализе стандартных образцов водки различающиеся результаты дают химические и хроматографические методики определения альдегидов и высших спиртов и т. д.

Выявив (обычно при визуальном исследовании) наличие двух или более максимумов на гистограмме результатов эксперимента, необходимо установить его причину. Это делается путем анализа сведений о процедуре анализа в лабораториях, которые запрашиваются у участников межлабораторного эксперимента. После выдвижения гипотезы об источнике различий между лабораториями (например, лаборатории применяют разные методики) их разбивают на соответствующие группы. Далее с использованием  $t$ -критерия Стьюдента или других способов сравнения средних значений в случае двух групп (см. разд. 2.5.3) или методом дисперсионного анализа в случае нескольких групп (см. разд. 2.6.2) проверяют гипотезу о том, что результаты измерений во всех лабораториях есть выборка из одной совокупности (см. гл. 2). Если такая гипотеза с принятым уровнем доверительной вероятности отвергается, приписывают стандартному образцу два или более средних содержания определяемого вещества, каждое из которых соответствует той или иной методике анализа. Такие опорные значения часто называю «методзависимыми».

**Пример.** При межлабораторном эксперименте по определению содержания общего холестерина в ведомственном стандартном образце сыворотки крови человека (участвовало 1200 лабораторий) была обнаружена полимодальность распределения результатов. Установлено, что это — результат применения разных методик анализа: прямого метода с использованием реактива Либермана — Бурхарда, экстракционного химического метода, трех различных ферментных методов. При однофакторном дисперсионном анализе (фактор — метод анализа) гипотеза о том, что все результаты есть случайная выборка из одной совокупности, была отвергнута при доверительной вероятности 0,999. Дальнейшая обработка данных проводилась после группировки лабораторий по используемым методам анализа, и стандартному образцу были приписаны несколько значений, каждое из которых соответствует одному из перечисленных методов анализа.

В данной ситуации единство измерений по объективным причинам не соблюдается.

Заметим, что при правильно организованном межлабораторном эксперименте в погрешность оценки опорного значения автоматически входит погрешность неоднородности СО.

Подробнее о межлабораторных экспериментах см. разд. 3.6.

4. Наконец, иногда аттестованное значение устанавливают, анализируя стандартный образец *в специально отобранных (selected) лабораториях*<sup>48</sup>. Фактически это тот же межлабораторный эксперимент, при организации которого стремятся не к максимально возможному числу лабораторий-участниц, а к более высокому их уровню. При этом число лабораторий невелико (обычно не более 10). Такой подход может быть эффективен, если организаторы эксперимента хорошо знают возможности лабораторий и могут выработать объективные критерии для отбора лабораторий – участниц эксперимента. К сожалению, выбор таких критериев – трудно формализуемая задача, и отбор лабораторий зачастую основывается на субъективных оценках, что может отразиться на результатах эксперимента.

### 3.2.5. Адекватность стандартных образцов

Одним из важнейших требований, предъявляемых к стандартным образцам сложного состава, является требование адекватности их исследуемым пробам. Погрешность, связанная с неадекватностью, определяется, с одной стороны, наличием скрытых систематических погрешностей метода анализа (т. е. зависимостью погрешности анализа от состава и свойств исследуемых проб) и, с другой стороны, различиями в составе рутинных проб и стандартного образца. Таким образом, говорить об адекватности стандартного образца вообще бессмысленно – можно говорить только об адекватности стандартного образца по отношению к конкретным пробам при использовании конкретной аналитической методики (методу анализа). Рассмотрим определение адекватности и связанные с ней погрешности, следуя [61, 62].

Предположим, что исследуемый СО предназначен для контроля точности некоторой методики КХА (в свою очередь предназначенной для анализа некоторого круга рутинных проб) и для этой методики справедлива зависимость

$$\Delta_c = f(C_1, C_2, \dots, C_p, \dots, K_1, K_2, \dots, K_m, \dots), \quad (3.18)$$

где  $\Delta_c$  – скрытая (не выявленная) систематическая погрешность анализа,  $C_1, C_2, \dots, C_p, \dots$  – содержание мешающих компонентов в исследуемой пробе (то есть присутствующих в анализируемых пробах веществ, отличных от определяемого и изменяющих величину аналитического сигнала),  $K_1, K_2, \dots, K_m, \dots$  – значения физико-химических показателей пробы, влияющих на результат анализа (такие, как показатели гранулометрического состава порошкообразных проб при спектральном анализе без растворения пробы)<sup>49</sup>. Предполагается, что при анализе с использованием рассматриваемой методики величины  $C_p, K_m$  не определяют (в противном случае погрешность не является скрытой и влияние величин  $C_p, K_m$  на результаты анализа можно учесть).

Вначале рассмотрим *адекватность стандартного образца одной рутинной пробе*<sup>50</sup>.

<sup>48</sup> Такие лаборатории называют также «экспертными», «образцовыми» и т. д.

<sup>49</sup> Мы не рассматриваем очевидный случай, когда зависимость вида (1) отсутствует. В такой ситуации говорить о неадекватности не приходится.

<sup>50</sup> Аналогично можно рассматривать адекватность двух стандартных образцов, двух рутинных проб и т. д.



Обозначим величины  $C_p$ ,  $K_m$  для СО как  $C_{CO,p}$ ,  $K_m$ , а для рутинной пробы как  $C_{рут,p}$ ,  $K_{рут,m}$ . Тогда неадекватность СО и рутинной пробы может быть охарактеризована величиной  $\Delta_{[CO; рут]}$ , которую можно назвать мерой адекватности этих образцов

$$\Delta_{[CO; рут]} = |f(C_{CO,p}, K_{CO,m} - C_{рут,p}, K_{рут,m})| = |\Delta_{CO} - \Delta_{рут}|. \quad (3.19)$$

Здесь  $\Delta_{CO}$  и  $\Delta_{рут}$  — систематические погрешности СО и рутинной пробы, определяемые матричными эффектами. Стандартный и рутинный образцы можно считать адекватными в случае, когда величина  $\Delta_{[CO; рут]}$  пренебрежимо мала по сравнению с допустимой (заданной) погрешностью методики анализа  $\Delta_d^{51}$  (см. разд. 3.2.1). Согласно принципу пренебрежения малыми погрешностями (разд. 2.6.1) для этого необходимо выполнение неравенства

$$\Delta_{[CO; рут]} < 1/3 \Delta_d, \quad (3.20)$$

где  $\Delta_d$  — допустимая основная погрешность анализа.

В случае невыполнения условия (3.20) СО и рутинную пробу нельзя считать адекватными друг другу (при использовании рассматриваемой методики анализа). Можно говорить также о более или менее адекватных образцах в зависимости от величины отношения  $\Delta_{[CO; рут]}/\Delta_d$ . Аналогичным образом можно рассмотреть адекватность любых двух проб.

К сожалению, для большинства методик химического анализа зависимость вида (3.18) неизвестна, и для ее нахождения необходимы специальные дорогостоящие исследования проб, для которых: а) содержание определяемого вещества известно с погрешностью, пренебрежимо малой по сравнению с погрешностью изучаемой методики, и б) известны параметры  $C_p$ ,  $K_m$ . Если необходимые пробы можно легко создать (например, задавать концентрации  $C_p$ , смешивая нужные компоненты), эксперимент можно планировать. В противном случае приходится исследовать большие совокупности проб и статистически обрабатывать полученные результаты; пример такого эксперимента описан в разд. 3.1.2.

Оценка адекватности двух образцов может оказаться необходимой, например, при сравнении проб, которые планируется систематически анализировать данным способом (стандартные образцы, пробы для внутрилабораторного контроля), при оценке возможности применения методики для анализа различных видов проб (в таких ситуациях берут два типичных образца разных видов) и т. д. В случае, когда зависимость (3.18) известна, нужно измерить величины  $C_p$ ,  $K_m$  для обеих проб, рассчитать и оценить  $\Delta_{[CO; рут]}$  по формулам (3.19, 3.20). Другой, более простой, способ оценки неадекватности — измерение величин  $\Delta_{CO}$  и  $\Delta_{рут}$  с использованием образцовой методики или методом «введено-найдено». К сожалению, в большинстве случаев это невозможно из-за отсутствия образцовой методики или невыполнения предположений, положенных в основу метода «введено-найдено» (см. разд. 3.4.4).

<sup>51</sup> Предполагается, что общая погрешность анализа не зависит от концентрации определяемого вещества. Если это не так, в неравенство (3) надо подставить погрешность анализа, соответствующую концентрации определяемого вещества в пробах 1 и 2, а если она различна, то наименьшую из двух.

Особенно обидно, что это справедливо именно для случаев, когда неадекватность велика (анализ твердых проб сложного состава).

*Адекватность стандартного образца и совокупности рутинных проб* требуется от всех образцов сравнения и СО (здесь и ниже говорится о стандартных образцах, однако все сказанное справедливо и для не узаконенных образцов сравнения). При этом подразумевается, как это обычно и бывает, что рассматриваемый стандартный образец предназначен для оценки правильности либо градуировки при анализе рутинных проб, относящихся к определенной совокупности.

**Примеры:** стандартный образец NIST #1710 для оценки правильности определения кадмия в алюминиевых сплавах; стандартный образец лиофилизированной мочи NIST # 2670 для оценки правильности определения токсичных металлов в моче человека и т. д.).

Также как в случае двух образцов, говорить об адекватности имеет смысл только применительно к конкретной методике химического анализа.

Предположим, как и раньше, что для данной методики химического анализа справедлива зависимость (3.18). Тогда составляющая систематической погрешности, определяемая матричными эффектами, для рассматриваемого стандартного образца равна

$$\Delta_{CO} = f(C_{CO,p}, K_{CO,m}), \quad (3.21)$$

Наряду с погрешностями аттестации, неоднородности и т. д. величина  $\Delta_{CO}$  определяет общую погрешность стандартного образца.

Введем аналогичные обозначения для рутинных проб (то есть тех проб, для проверки правильности анализа которых предназначен рассматриваемый стандартный образец):

$$\Delta_i = f(C_{i,p}, K_{i,m}), \quad (3.22)$$

где  $\Delta_i$  — скрытая систематическая погрешность для  $i$ -й рутинной пробы, определяемая неизвестными содержанием мешающих компонентов  $C_{i,p}$  и значениями физико-химических показателей пробы, влияющих на результат анализа  $K_{i,m}$ , в этой пробе.

На рис. 3.3 условно показаны распределение величин  $\Delta_i$  для рассматриваемой совокупности рутинных проб, и величина  $\Delta_{CO}$ . Будем, как обычно, полагать рассматриваемое распределение величин  $\Delta_i$  нормальным. Тогда ~95% всех  $\Delta_i$  лежат в интервале  $\{\Delta_L, \Delta_H\}$ , где  $\Delta_L = \bar{\Delta}_i - 2s(\Delta_i)$ ,  $\Delta_H = \bar{\Delta}_i + 2s(\Delta_i)$ ,  $\bar{\Delta}_i = \sum_{i=1}^I \Delta_i / I$ ,  $I$  — число

рутинных проб,  $s(\Delta_i) = \sqrt{\frac{\sum_{i=1}^I (\Delta_i - \bar{\Delta}_i)^2}{I-1}}$  — стандартное отклонение всех  $\Delta_i$ .

В качестве меры адекватности стандартного образца естественно принять величину МА, представляющую собой максимальную (по абсолютной величине) разность из двух:

$$MA = \max\{|\Delta_L - \Delta_{CO}|, |\Delta_H - \Delta_{CO}|\}. \quad (3.23)$$



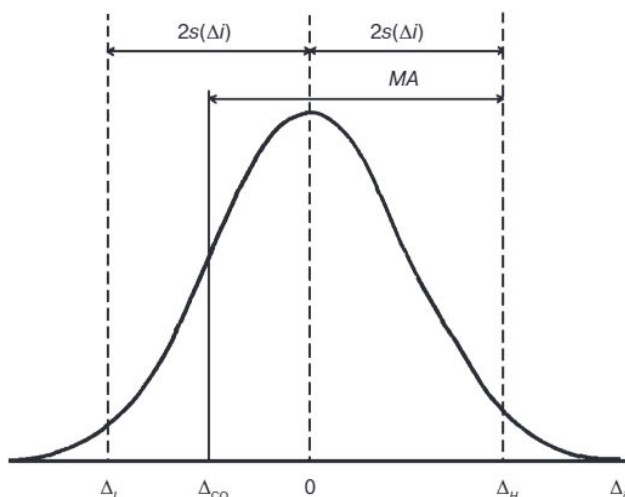


Рис. 3.3. Адекватность стандартного образца и совокупности рутинных образцов. Показано нормальное распределение величин  $\Delta_i$ , его  $\sim 95\%$  квантили ( $\Delta_L$ ,  $\Delta_H$ ), величина  $\Delta_{CO}$  и мера адекватности  $MA = \max\{|\Delta_L - \Delta_{CO}|, |\Delta_H - \Delta_{CO}|\}$

Другими словами, это — максимальное из расстояний величины систематической погрешности стандартного образца до одного из двух квантилей распределения систематических погрешностей рутинных проб  $Q(0,025)$ ;  $Q(0,975)$  (см. рис. 3.3). Аналогичный подход возможен и в случаях, когда распределение величин  $\Delta_i$  отклоняется от нормального.

Стандартный образец адекватен совокупности рутинных проб, если выполняется соотношение

$$MA < \frac{1}{3}\Delta_A. \quad (3.24)$$

Как указывалось выше, на практике вид зависимости (3.18) для конкретных методик анализа чаще всего неизвестен. Одно из немногих исключений: рутинные методики определения концентрации общего холестерина в сыворотке крови человека, результаты исследования которых изложены нами в Примере в разд. 3.1.2.

**Пример.** Как там указано, подставляя в эмпирические уравнения (А–Д) 5% квантили распределений мешающих компонентов (то есть считая распределения мешающих компонент независимыми), находим, что при неудачном сочетании мешающих компонентов величина систематической погрешности может составлять для методов (А–Д)  $(0 \div +14)$ ;  $(-6 \div +13)$ ;  $(-17 \div +1)$ ;  $(-19 \div +26)$  процентов соответственно. Фактически в скобках указаны величины  $\Delta_L$  и  $\Delta_H$  для каждого метода. Более того, наличие в уравнениях значительных свободных членов указывает на возможное существование неизвестных нам факторов, влияющих на результаты анализа.

Оптимальным является случай, когда величина  $\Delta_{CO}$  лежит в середине интервала  $\{\Delta_L, \Delta_H\}$ . В этом случае мера адекватности  $MA$  наименьшая и равна  $MA = (\Delta_L + \Delta_H)^2$ . Для рассмотренных нами методов определения концентрации общего холестерина наименьшее возможное значение  $MA$  равно 7, 9, 9 и 22% для методов (А–Д) соответственно. В то же время, согласно различным национальным и международным документам, общая относительная погрешность определения концентрации общего холестерина в сыворотке крови человека, не

должна превышать 3–6% — бóльшая погрешность не позволяет достоверно ставить диагноз обследуемым. Таким образом, поскольку ни для одного из рутинных методов соотношение (3.24) не выполняется, создание адекватного стандартного образца в рассмотренных в примере случаях невозможно.

Другой способ оценки величины  $MA$ , не требующий знания зависимости вида (3.18) и распределения значений  $C_p$ ,  $K_m$  в рутинных пробах, — оценка величин систематической погрешности  $\Delta_c$  для репрезентативной выборки из всей совокупности рутинных проб методом «введено-найдено» или с использованием метода сравнения. К сожалению, в наиболее важных случаях это практически невозможно. Помимо указанных выше причин (неприменимость метода «введено-найдено», отсутствие метода сравнения) часто оказывается сложным получение представительной выборки рутинных проб. Так, в рассмотренном примере имеются использованные нами для оценок литературные данные о распределении мешающих компонентов в различных популяциях (у жителей разных городов и стран), а получение представительной выборки проб сыворотки крови всех жителей Земли вряд ли реально.

Хотя исчерпывающие исследования, дающие оценку неадекватности стандартных образцов, в силу их сложности и дороговизны, редки, можно с уверенностью утверждать, что ситуация, когда создание адекватного стандартного образца невозможно, типична. Более того, общей тенденцией является увеличение числа упрощенных методик, методик с использованием сенсоров и тест-методов, для которых скрытые систематические погрешности заведомо велики и создание адекватных стандартных образцов невозможно. Конечно, это не относится к дорогим и сложным методам анализа, позволяющим добиться высокой правильности (для рассмотренного случая определения общего холестерина в сыворотке крови человека таким методом является масс-спектрометрия с изотопным разбавлением), а также случаев анализа простых по составу объектов.

Однако отсутствие адекватных стандартных образцов влечет за собой необходимость пересмотра существующей в настоящее время практики, основанной на пренебрежении систематическими погрешностями либо возможности их учета. В подобных ситуациях практически невозможно ни устранение систематических погрешностей в ходе градуировки, ни даже корректная оценка погрешностей анализа индивидуальных рутинных проб (исключение — популяционные исследования [54, 64]).

Как же действовать при отсутствии адекватных стандартных образцов? Возможны следующие варианты.

1. Неадекватность может учитываться непосредственно. Для этого при расчете общей погрешности  $CO$  по формуле (3.2) в качестве оценки  $\xi_{MA}$  включают величину  $s_{MA} = (1,4–1,7)MA$ . Это — стандартный прием замены систематической погрешности на случайную (см. также разд. 3.5.2.3.2). Тогда общая погрешность (неопределенность) оказывается гораздо больше, чем рассчитанная только с учетом случайных погрешностей, и может быть оценена либо по экспериментальным данным (например, как это сделано выше), либо с привлечением экспертных оценок. То же самое можно сделать, просто увеличив доверительный интервал для результата анализа. Однако простая констатация того, что реальные погрешности анализа могут оказаться намного больше, чем принято считать, мало что дает — требования к погреш-



ностям химического анализа чаще всего довольно жесткие (как в рассмотренном выше примере определения общего холестерина в сыворотке крови человека).

2. Для снижения погрешностей измерения концентрации возможно использование последовательной стратегии, состоящей в проведении дополнительных исследований (более точное определение того же показателя, определение других показателей) при получении результатов анализа, нуждающихся в уточнении. Например, можно дополнительно анализировать объекты окружающей среды в случаях, когда результат анализа лежит вблизи ПДК (см. также разд. 3.5.2.4.2).

3. Наконец, можно использовать нехимическую информацию при принятии решения по результатам анализа. Фактически это уже делается во многих случаях (например, на химических производствах, в медицине и т. д.) и, по сути дела, в значительной мере страхует от возможных последствий ошибочной интерпретации результатов химического анализа (в том числе отсутствия учета погрешности неадекватности).

Обсуждая адекватность СО, необходимо упомянуть и понятие их «коммутативности» [1, 65]. *Коммутативностью* называют свойство стандартного образца при определении какой-либо величины в нем по двум разным методикам демонстрировать результаты, соотношение между которыми близко к соотношению результатов, полученных по этим методикам для рутинных проб.

### 3.2.6. Утверждение (признание) стандартных образцов

После разработки технологии приготовления стандартного образца того или иного типа, по ней готовят стандартный образец, оценивают его стабильность и однородность, устанавливают способ оценки содержаний аналитов и их погрешностей (неопределенностей). Во многих случаях оценивают также адекватность стандартного образца при его использовании по назначению.

Результаты всех этих исследований излагаются в форме документа (отчета)<sup>52</sup> и передаются в организацию, которая должна их утверждать — если, конечно, стандартный образец не разрабатывался в самой организации. Наибольший интерес представляет утверждение международных и национальных стандартных образцов.

Международные стандартные образцы утверждаются международными организациями, а национальные — уполномоченным национальным органом (органами). Высоким авторитетом пользуются стандартные образцы, утвержденные Национальным институтом стандартов и технологий (NIST), США, LGC (Великобритания) и т. д. Национальным органом, выполняющим такую работу в России, является Росстандарт (методический центр — Уральский научно-исследовательский институт метрологии), в СНГ — Межгосударственный совет по стандартизации, метрологии и сертификации. Согласно закону [2] в России обязательному утверждению (узаконению) подлежат СО, используемые в сфере государственного регулирования единства измерений.

Утверждению предшествует экспертиза представленной документации и в некоторых случаях дополнительное экспериментальное исследование стандартных

<sup>52</sup> В России требуется также оформлять техническое задание на разработку образца с указанием требований, которым должен удовлетворять создаваемый стандартный образец [52].

образцов. При этом обычно учитывается не только уровень разработки, но и авторитет организации-разработчика. После утверждения стандартные образцы заносятся в соответствующие реестры.

*Стандартные образцы, утверждаемые на уровне организаций, корпораций, объединений, ведомств и других юридических лиц – СОГНМИ, ОСО, СОП – разрабатываются, утверждаются и производятся этими юридическими лицами самостоятельно.*

### 3.3. Метрологические характеристики методик КХА

Прежде всего отметим, что методика количественного химического анализа всегда предназначается для измерения концентрации *определенных веществ (аналитов) в определенных объектах*.

Ввиду сложности и потенциальной нестабильности процедур химического анализа применяемая в других областях измерений классификация погрешностей (см. разд. 1.3.5) здесь оказывается недостаточной. Для метрологической оценки качества методик количественного химического анализа наряду с обычными метрологическими характеристиками измерений используют специфические показатели.

Основными метрологическими характеристиками методик КХА являются характеристики случайных и систематических погрешностей (показатели прецизионности и правильности, см. разд. 3.3.1). Там, где возможно, мы будем ориентироваться на международный документ [27], в котором даны терминология и некоторые способы оценки этих характеристик<sup>53</sup>.

#### 3.3.1. Терминология для прецизионности и ее составляющих

В [27] для важнейшей характеристики качества результатов измерений – случайной погрешности – введен общий термин *прецизионность* (англ. precision) – *степень близости друг к другу независимых результатов измерений, полученных в конкретных регламентированных условиях*.

Для удобства дальнейшего изложения определим термин «серия».

*Серия (англ. run) это совокупность образцов, исследуемых одним и тем же методом, в одной и той же лаборатории, при одной и той же калибровке, одним и тем же оператором, с использованием одного и того же оборудования, в пределах короткого промежутка времени*<sup>54</sup>. В качестве синонимов мы будем использовать также обороты «аналитическая серия» и «серия измерений»<sup>55</sup>.

Как метрологические характеристики методик КХА, рассматриваются следующие компоненты *прецизионности*.

<sup>53</sup> Вообще говоря, в документе ИСО [27] не говорится, к каким именно измерениям относится его содержание (вероятно, к любым). Однако все примеры взяты из химического анализа.

<sup>54</sup> В [27] для результатов, полученных в одной серии, использован оборот «результаты, полученные в условиях повторяемости (сходимости)».

<sup>55</sup> В специальных случаях под «серией» может пониматься и совокупность проб, проанализированных в течение определенного промежутка времени, но разными лаборантами (или на разных приборах), например, в течение одной смены. Такие ситуации мы будем оговаривать особо.



**Повторяемость (сходимость)**<sup>56</sup> (англ. repeatability) — это прецизионность в условиях повторяемости (сходимости), при которых независимые результаты измерений получаются одним и тем же методом на идентичных объектах, в одной и той же лаборатории, одним и тем же оператором, с использованием одного и того же оборудования, в пределах короткого промежутка времени. Другими словами, *повторяемость (сходимость) характеризует разброс результатов в пределах одной аналитической серии. Мерой повторяемости (сходимости) является стандартное отклонение повторяемости (сходимости)  $s_r$  либо предел повторяемости (сходимости) (repeatability limit)  $r \approx 2,77s_r$ .*

**Воспроизводимость** (reproducibility) — это прецизионность в условиях воспроизводимости, при которых результаты измерений получают одним и тем же методом, на идентичных объектах испытаний, но в разных лабораториях, разными операторами, с использованием различного оборудования. То есть *воспроизводимость характеризует разброс результатов в разных лабораториях. Мерой воспроизводимости является стандартное отклонение результатов измерений, полученных в условиях воспроизводимости  $s_R$ , либо предел воспроизводимости (reproducibility limit)  $R \approx 2,77s_r$ .*

Промежуточная прецизионность (intermediate precision) характеризующая прецизионность результатов измерений, полученных *в одной лаборатории, но в разных сериях*, то есть при изменяющихся факторах — времени ( $T$ ), калибровке ( $C$ ), оборудовании ( $I$ ), операторе ( $L$ ), причем изменяться может либо один из факторов, либо несколько в разных сочетаниях. Мерой промежуточной прецизионности является соответствующее стандартное отклонение  $s_{IPTCL}$ , причем в скобках указываются изменяющиеся факторы. Промежуточную прецизионность при всех четырех изменяющихся факторах иногда называют *внутрилабораторной прецизионностью*.

При исследовании методик КХА оценивают повторяемость (сходимость) и, когда это удастся, воспроизводимость (см. ниже). Промежуточную прецизионность при этом практически никогда не оценивают. Причина в том, что, как следует из определения, промежуточных прецизионностей может быть несколько — для каждого набора изменяющихся от серии к серии факторов своя. Оценивать каждую из них было бы чрезвычайно дорого и трудоемко.

Отдельно следует рассмотреть вопрос об обозначении  $s_r$ . Во многих работах это не стандартное отклонение повторяемости (сходимости), а относительное стандартное отклонение  $s/X$ , где  $s$  — какое-либо стандартное отклонение, а  $\bar{X}$  — среднее значение результатов измерений (см. разд. 2.2). Во избежание путаницы мы, как это делается во многих англоязычных публикациях, для относительного стандартного отклонения используем аббревиатуру **RSD** (от *relative standard deviation*) с указанием конкретной составляющей прецизионности, если из контекста неясно, о какой из них идет речь (например, **RSD<sub>r</sub>** — *относительное стандартное отклонение повторяемости (сходимости)*, **RSD<sub>R</sub>** — *относительное стандартное отклонение воспроизводимости*).

Рассмотренные компоненты прецизионности и общая случайная погрешность методики связаны между собой.

<sup>56</sup> Термины «повторяемость» и «сходимость» являются синонимами.

### 3.3.2. Терминология для правильности КХА

*Правильность* (trueness) *есть степень близости среднего значения, полученного на основании большой серии результатов измерений (или результатов испытаний), к принятому опорному значению*, то есть это характеристика систематической погрешности. Количественно правильность может быть охарактеризована *систематической погрешностью*  $\Delta_c$

$$\Delta_c = \mu - \tau, \quad (3.25)$$

где  $\mu$  — математическое ожидание результатов измерений,  $\tau$  — истинное значение измеряемой величины. Понятно, что ни  $\mu$ , ни  $\tau$ , строго говоря, неизвестны — мы имеем лишь их более или менее точные их оценки. Поэтому на практике используют *смещение*  $B$ :

$$B = \bar{X} - X_0. \quad (3.26)$$

Здесь  $\bar{X}$  — среднее значение результатов нескольких повторных измерений образца сравнения (стандартного образца), полученных в одинаковых условиях,  $X_0$  — принятое опорное (аттестованное) значение для этого образца (см. разд. 3.2). Часто используют относительное смещение, выраженное в процентах,  $\%B$ :

$$\%B = \frac{\bar{X} - X_0}{X_0} \times 100\%. \quad (3.27)$$

Как и в «классической» метрологии, для результата единичного измерения  $X_i$  используется величина погрешности результата измерения  $e_i$ :

$$e_i = X_i - \tau, \quad (3.28)$$

аналогичная абсолютной погрешности измерительного прибора  $\Delta$  (см. разд. 1.4.1). Как и для измерительных приборов, используют и относительную погрешность результата измерения  $e_r = e_i/\tau$ , которую часто выражают в процентах.

Заметим, что при представлении результатов количественного химического анализа термин «погрешность» применяют в несколько ином смысле, понимая под ним общую, или суммарную, погрешность результата анализа. При этом во многих случаях пренебрегают скрытыми систематическими погрешностями и погрешность указывают как половину доверительного интервала. В этом случае результат анализа записывают в форме  $(X \pm \Delta X)$ , где  $X$  — окончательно приводимое значение,  $\Delta X$  — «погрешность анализа», с указанием принятого уровня доверительной вероятности. Способ оценки расчета доверительного интервала зависит от схемы проведения измерений.

Для тех же целей используется понятие «неопределенность». Подробно об использовании понятий «погрешность» и «неопределенность» при представлении результатов КХА рассказано в разд. 3.5.2 и сл.



### 3.3.3. Другие метрологические характеристики методик КХА

**Градуировочная характеристика и чувствительность** методики КХА определяются так же, как для средств измерений (см. разд. 1.4.1) с тем отличием, что здесь это зависимость *аналитического сигнала* (см. разд. 3.4.1) от измеряемой величины (*концентрации*). В случае нелинейной градуировочной характеристики часто выделяют ее линейную область, т. е. область, в пределах которой градуировочная характеристика практически линейна.

Важной характеристикой методики анализа является **предел обнаружения** — минимальное значение концентрации, которое может быть с уверенностью обнаружено. Другое определение и способы оценки предела обнаружения даны в разд. 3.4.2.

**Робастность** — свойство методики анализа, характеризующее ее устойчивость к небольшим изменениям условий измерений. Это качественное понятие. Как известно, на результаты анализа могут влиять многие факторы, как контролируемые, так и неконтролируемые в ходе измерений. Обычно робастной считается методика, для которой ни один из таких факторов не имеет значимо большего (по сравнению с другими) влияния на результаты анализа. Заметим, что понятие робастности частично пересекается с понятиями нормальных и рабочих условий эксплуатации средств измерений.

**Селективность и специфичность** — средства выражения степени влияния веществ матрицы пробы на результаты анализа по данной методике. Предложена [66] количественная интерпретация этих понятий, требующая знания всех компонентов, из которых состоит анализируемое вещество, и зависимости результата анализа от концентрации каждого из них и их сочетаний. Поскольку полная информация такого рода практически недоступна, понятия селективности и специфичности чаще всего используют как качественные. Так, хромато-масс-спектрометрия в анализе органических проб сложного состава более селективна, чем хроматография — один пик на хроматограмме может соответствовать нескольким веществам, каждое из которых дает свой пик (пики) в масс-спектрометре. Подробнее об этих характеристиках см. разд. 3.3.5.

В качестве метрологических характеристик аналитических методик используются и другие, менее универсальные показатели, например, минимальная площадь анализируемого объекта, вносящая вклад в аналитический сигнал для локальных неразрушающих методов анализа.

Основные способы оценки метрологических характеристик КХА описаны ниже.

## 3.4. Исследование методик количественного химического анализа

Во многих случаях, после того как методика КХА реализована, ее исследуют, т. е. оценивают характеризующие ее метрологические характеристики (см. разд. 3.3). Это делают всегда для методик, подлежащих стандартизации.

В этой области существует терминологический разнобой. Широко распространен термин *«аттестация методик»*. Под этим подразумевают либо только утверждение (узаконение) методики, либо (в совокупности) исследование метро-

логических характеристик и последующее узаконение методики. В последнее время используют также аналогичный термин «*валидация*».

Мы будем называть *аттестацией* лишь стадию утверждения (узаконения) методики, выделяя ее исследование в отдельную стадию. Это связано с тем, что методика, как правило, исследуется в компетентных лабораториях опытным путем, зачастую на протяжении длительного времени, в то время как узаконение (собственно аттестация) проводится специалистами по экспертизе документов в аккредитованных организациях и не предусматривает проведения экспериментов. Она рассмотрена в разд. 3.4.

Рассмотрим основные способы исследования методик КХА для оценки их метрологических характеристик.

### 3.4.1. Градуировочная характеристика и градуировка

В количественном химическом анализе как варианте косвенных измерений значение измеряемой величины — концентрации  $C$  — находят на основании зависимости между этой величиной и другими величинами, поддающимися прямым измерениям ( $X_1, X_2, \dots, X_n$ ). Для методики анализа градуировочная характеристика (статистическая характеристика преобразования) есть зависимость (выходного) аналитического сигнала  $A$  от (входного) — концентрации  $C$

$$A = f(C). \quad (3.29)$$

Поскольку химический анализ представляет собой косвенное измерение (см. гл. 1), аналитический сигнал  $A$ , в свою очередь, представляет собой некоторую функцию величин, поддающихся прямым измерениям  $X_1, X_2, \dots, X_n$  и в ряде случаев, времени  $t$

$$A = f(X_1, X_2, \dots, X_n, t). \quad (3.30)$$

Выбор вида функции (3.29) для каждого метода анализа определяется как принципом, положенным в основу метода, так и соображениями удобства. Конечно, стремятся работать в области, где эта зависимость линейная. Почти всегда аналитический сигнал рассчитывается как разность двух величин.

**Примеры.** В фотометрии и пламенном атомно-абсорбционном анализе в качестве аналитического сигнала чаще всего используют оптическую плотность — разность логарифмов интенсивностей исходного и прошедшего через поглощающую среду света с определенной длиной волны; при выполнении определенных условий оптическая плотность пропорциональна концентрации определяемого вещества в поглощающем слое. При объемном анализе по методу обратного титрования это разность между объемом титранта в бюретке до и после титрования. В атомно-эмиссионной спектроскопии аналитический сигнал есть интеграл зависимости интенсивности испускаемого атомами или молекулами света от времени за вычетом фонового сигнала и т. д.

Зависимость (3.29) в количественном химическом анализе называют *градуировочной характеристикой*.

Для нахождения вида градуировочной характеристики, т. е. для проведения градуировки, находят аналитический сигнал для нескольких образцов сравнения



с разными концентрациями определяемого вещества (ниже – *градуировочные образцы*). Измерения обычно выполняются в одной аналитической серии. Концентрации определяемого вещества в градуировочных образцах выбираются такими, чтобы перекрыть весь диапазон, для измерений в котором предназначена методика. При исследовании методики (в отличие от последующего ее рутинного применения) анализируют достаточно много градуировочных образцов, проводя для каждого несколько измерений; оптимально использовать 6–9 образцов, анализировать более 12–14 нет смысла. В большинстве случаев градуировку проводят по градуировочным образцам с равноотстоящими концентрациями («эквидистантные»). С метрологической точки зрения такой план эксперимента оптимален (максимально информативен) при выполнении стандартных предположений о нормальном распределении погрешностей и однородности дисперсий. В случае неоднородности дисперсий (гетероскедастичность) эквидистантное распределение градуировочных образцов по концентрации, вообще говоря, не оптимально, однако для практических целей вполне удовлетворительно.

Далее решают классическую задачу регрессионного анализа (разд. 2.7.2). Отметим особенности применения регрессионного анализа при изучении градуировочных характеристик количественного химического анализа.

Очень часто при визуальном исследовании оказывается, что градуировочная характеристика практически линейна в достаточно широком диапазоне концентраций, но отклоняется от линейности при высоких или (и) низких концентрациях. В таких случаях часто ограничиваются линейной областью градуировочной характеристики, «подстраивая» методику так, чтобы аналитический сигнал всегда попадал в область линейности. Этого можно достичь разбавлением проб либо их растворов, изменяя соотношение реагентов и т. д.

При работе в диапазоне концентраций более 1.5–2 порядков часто возникает неоднородность дисперсий, которую необходимо учитывать при расчетах, например, введением весового фактора (см. разд. 2.7.2.3).

Если градуировочные образцы представляют собой растворы чистых веществ, погрешностью их аттестации (т. е. погрешностью независимой переменной) можно пренебречь. Действительно, при выполнении правил приготовления растворов (используя достаточно чистое вещество, точные весы и мерные колбы, термостатируя растворитель и т. д.) легко добиться относительной погрешности концентрации градуировочных образцов не более 0.5%, что пренебрежимо мало по сравнению с погрешностью большинства методов химического анализа.

Однако такое предположение некорректно для образцов сравнения сложного состава, аттестуемых методом межлабораторного эксперимента (см. выше), а также для образцов, опорное значение которых найдено методом добавок. В таких случаях при регрессионном анализе необходимо учитывать погрешность независимой переменной (см. разд. 2.7.2.3).

Чувствительность методики количественного химического анализа определяется аналогично чувствительности средств измерений – как производная градуировочной характеристики  $f'(C) = dA/dC$ . В случае линейной градуировочной характеристики  $A = KC$  чувствительность равна  $K$ .

Заметим, что чувствительность не указывает непосредственно минимальную концентрацию, которую можно достоверно определить с помощью данной методики. Эту величину характеризует предел обнаружения, связанный как с чувствительностью, так и с прецизионностью методики (см. ниже, п. 3.3.2).

В практике рутинного анализа градуировку обычно выполняют достаточно часто, иногда даже в каждой аналитической серии. Однако иногда на основе опыта полагают, что вид градуировочной характеристики, т. е. вид функции (3.23), неизменен и от одной аналитической серии к другой изменяются лишь ее параметры, например, для линейной градуировочной характеристики величина  $K$ . При этом для каждой аналитической серии лишь уточняют эти параметры (анализируя один или два градуировочных образца), а полную градуировочную характеристику строят лишь изредка («контроль стабильности градуировочной характеристики»). После построения (или уточнения параметров) градуировочной характеристики измеряют аналитический сигнал исследуемых проб и по нему рассчитывают концентрацию определяемого вещества.

### 3.4.2. Предел обнаружения

Обычно *предел обнаружения* определяется как минимальное значение концентрации, которое может быть с уверенностью обнаружено (с вероятностью  $P$ ) при использовании минимального статистически значимого сигнала как порогового значения. При этом пороговое значение понимается как минимальное статистически значимое значение «чистого» сигнала, который может быть зарегистрирован с вероятностью  $P$ . Далее мы будем использовать это определение.

Рассмотрим последовательность действий по оценке предела обнаружения аналитической методики для случая линейной градуировочной характеристики вида

$$A = \alpha + \beta C \quad (3.31)$$

и выполнения основных предположений метода наименьших квадратов (см. разд. 2.7.2.2).

1. С использованием одной аналитической серии проводят измерения, обычные для построения градуировочной характеристики для данной методики.

2. Находят оценки  $a$  и  $b$  коэффициентов  $\alpha$ ,  $\beta$ , а также  $s_a$ ,  $s_b$  и  $s_0$  (см. разд. 2.7.2.2). Кроме того, рассчитывают коэффициент линейной корреляции  $r(A, C)$  между аналитическим сигналом и концентрацией (см. разд. 2.7.1).

3. Определяют минимально значимый «чистый» аналитический сигнал  $S_c$  по формуле

$$S_c = t(\bar{P}, f) s_{ac}, \quad (3.32)$$

где  $t(\bar{P}, f)$  — процентная точка  $t$ -распределения Стьюдента, соответствующая выбранной доверительной вероятности  $\bar{P}$  (чаще всего принимают  $\bar{P} = 0,99$  или близкую к ней), и числу степеней свободы  $f$ , равному числу измерений минус 1,  $s_{ac}$  — стандартное отклонение чистого аналитического сигнала, равное



$$s_{ac} = \sqrt{s_a^2 + s_0^2}. \quad (3.33)$$

Собственно предел обнаружения определяемого вещества  $C_D$  рассчитывают по формуле

$$C_D = 2 \frac{S_c}{b} \frac{(1 + r(A, C)(s_a / s_{ac})(s_b / b)t(\bar{P}, f))}{1 - t(\bar{P}, f)(s_b / b)^2}. \quad (3.34)$$

В случае нелинейной градуировочной характеристики проводят измерения с использованием нескольких образцов сравнения с концентрациями, близкими к ожидаемому пределу обнаружения, и по результатам этих измерений оценивают предел обнаружения по формуле (3.34). Практически любая зависимость в узком интервале измеряемых величин неплохо описывается линейной зависимостью. Кроме того, не возникает неоднородность дисперсии, усложняющая расчеты.

Запись найденного для методики предела обнаружения должна обязательно сопровождаться указанием доверительной вероятности.

Используют и другой, более простой способ нахождения предела обнаружения  $C_D$ :

$$C_D = k s_0, \quad (3.35)$$

где  $s_0$  — стандартное отклонение результатов измерений при концентрации, близкой к пределу обнаружения;  $k$  — коэффициент, определяющий доверительную вероятность. Если число измерений при оценке  $s_0$  достаточно велико (8–10 или более), то  $k = 2$  соответствует  $\bar{P} \approx 0,95$ ,  $k = 3$  соответствует  $\bar{P} \approx 0,998$  и т. д. в предположении о нормальном распределении результатов. При этом часто используют следующую форму записи: «предел обнаружения [аналита по данной методике], рассчитанный по 3s-критерию, составляет...» при  $k = 2$  (или ... при  $k = 3$ )».

В таком виде способ применим в отсутствие поправки контрольного («холодного») опыта. Если такая поправка существенна, то есть аналитический сигнал достоверно отличается от нуля в отсутствие аналита, то необходимо учитывать погрешность оценки этой поправки.

В этом случае концентрация аналита рассчитывается как разность результата измерений концентрации в пробе и концентрации, найденной в контрольном опыте. Тогда в соответствии с правилами сложения дисперсий (разд. 2.6.1) дисперсия погрешности измерения концентрации аналита  $s_0^2$  равна сумме дисперсий результатов измерения суммарной концентрации в пробе  $s_{\text{пр}}^2$  и дисперсии оценки концентрации для контрольного опыта  $s_{\text{к.о.}}^2$ :

$$s_0^2 \cong s_{\text{пр}}^2 + s_{\text{к.о.}}^2.$$

Полагая для простоты, что измерение поправки контрольного опыта выполнялось так же, как измерение для пробы, и  $s_{\text{пр}}^2 \approx s_{\text{к.о.}}^2 = s^2$  (то есть постоянство погрешности в некотором интервале концентраций), получаем  $s_0 \cong \sqrt{2s^2}$ . Эту величину и надо подставлять в формулу (3.35). Для большей точности можно учитывать отличия погрешностей измерений при разных концентрациях, а также возможную асимметрию распределения вблизи предела обнаружения, но чаще всего это не имеет особого смысла.

*Пределы обнаружения для какой-либо методики, рассчитанные обоими вышеописанными способами, не слишком отличаются друг от друга.*

При исследовании методик предел обнаружения оценивают не всегда. Для многих областей анализа концентрация аналитов в объектах, для исследования которых предназначена методика, заведомо намного больше предела обнаружения известных методик. Такова ситуация с определением ряда веществ, таких как глюкозы или холестерина, в сыворотке крови человека. В подобных случаях ограничиваются исследованием погрешностей анализа в диапазоне концентраций, в котором могут лежать концентрации в реальных объектах, для исследования которых предназначена методика.

### 3.4.3. Оценка прецизионности

При оценке метрологических характеристик методик в ходе их аттестации (валидации) оценивают повторяемость и воспроизводимость. Их характеристики и приводят в тексте методик. Промежуточную прецизионность при этом не определяют, что связано с тем, что  $s_{lp}$  различно для разных сочетаний меняющихся факторов (см. разд. 3.3.1). Фактически для методики можно было бы найти несколько разных  $s_{lp}$ , слишком дорого и не всегда возможно. Зато промежуточная прецизионность чрезвычайно важна при контроле стабильности (см. разд. 4.5.1).

*Оценка повторяемости (сходимости) в одной аналитической серии.*

Если требуется оценить только повторяемость (сходимость), это можно сделать, выполнив  $N$  (не менее 8–10) измерений одного и того же однородного материала в одной серии. Для полученных результатов  $X_i$  рассчитывают (после отброса грубых промахов, см. разд. 2.4.2) стандартное отклонение повторяемости (сходимости)  $s_r$ :

$$s_r = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1}}, \text{ где } \bar{X} = \frac{\sum_{i=1}^N X_i}{N}. \quad (3.36)$$

В случае, когда это делается в ходе оценки метрологических характеристик валидации методики в конкретной лаборатории, полученное значение сравнивают с нормативом (см. разд. 4.5.1.2).

Совместная оценка повторяемости (сходимости) и воспроизводимости методики.

Как было сказано в разд. 3.4, при исследовании методик КХА, подлежащих стандартизации, обычно оценивают повторяемость (сходимость) и воспроизводимость методики. Для совместной оценки соответствующих стандартных отклонений проводят межлабораторный эксперимент, в рамках которого в достаточно большое (не менее 8) число лабораторий рассылают комплект из нескольких (обычно не менее 5) контрольных материалов с разной концентрацией аналита<sup>57</sup>. Для каждого контрольного материала в каждой лаборатории выполняется несколько измерений в одной аналитической серии. Фактически для каждого материала выполняется эксперимент, описанный в разд. 2.6.2 (см. рис. 2.7, табл. 2.3 и сл.). Соответственно,

<sup>57</sup> В [67] эти материалы названы «образцами для оценивания (ОО).



и обработка данных проводится в рамках модели (2.45) методом дисперсионного анализа.

В отсутствие априорной информации о соотношении разных видов погрешности точные рекомендации по выбору плана эксперимента дать нельзя. Общие же рекомендации таковы. Следует стремиться к возможно большему числу лабораторий-участниц эксперимента (нежелательно, чтобы их число было меньше 8, оптимальное количество 20–30). Число измерений концентрации одного аналита в каждой аналитической серии обычно выбирают в пределах от 2 до 4; увеличение числа измерений не дает существенного выигрыша и резко повышает стоимость эксперимента.

Результаты измерений рассматривают с точки зрения наличия подозрительно выделяющихся значений. При этом грубые промахи («выбросы») выявляют как в каждой лаборатории, так и по средним значениям для лабораторий (см. разд. 2.4.3, 2.5.2). Отбрасывают как результаты единичных измерений в конкретной лаборатории, так и все результаты для «плохой» лаборатории.

Далее результаты для каждого материала обрабатывают методом однофакторного дисперсионного анализа в рамках модели (2.45). Как и в разд. 2.6.2 будем полагать, что случайные эффекты  $\alpha_j$  и  $e_{ij}$  распределены нормально:  $\alpha_j \in N(0, \sigma_\alpha^2)$ ,  $e_{ij} \in N(0, \sigma_e^2)$ , а объемы всех выборок (число измерений в каждой лаборатории) одинаковы, т. е.  $N_1 = N_2 = \dots = N_m = N$ .

Проверяют гипотезу об отсутствии межлабораторного разброса  $H_0: (\alpha_j = 0)$ ; способ проверки дан в разд. 2.6.2. Почти всегда она отвергается с принятым уровнем доверительной вероятности (чаще всего 0,95).

Оценки стандартных отклонений повторяемости (сходимости)  $s_r$  и воспроизводимости  $s_R$  рассчитываются как

$$s_r = \sqrt{s_{\text{ост}}^2}, \quad s_R = \sqrt{\frac{\sum_{j=1}^m (\bar{X}_j - \bar{\bar{X}})^2}{m-1} + \left(1 - \frac{1}{N}\right) s_r^2}. \quad (3.37)$$

Здесь использованы обозначения из табл. 2.3 и 2.4.

Для получения полной характеристики случайных погрешностей методики необходимо знать зависимость их от концентрации определяемого вещества. Для этого проводят измерения, описанные выше, с использованием нескольких контрольных образцов с разной концентрацией аналита. По результатам анализа строят график концентрационной зависимости компонентов случайной погрешности. Обычно такие графики строят в координатах «относительное стандартное отклонение — концентрация» (рис. 3.4).

Другими словами, для каждого контрольного материала находят величины  $RSD_r = s_r / \bar{\bar{X}}$  и  $RSD_R = s_R / \bar{\bar{X}}$ , после чего строят зависимости этих величин от концентрации. Часто они имеют U-образный вид, причем обнаруживается достаточно большой (не меньше порядка величины) интервал концентраций с постоянными относительными стандартными отклонениями, которые возрастают вблизи предела обнаружения и в области больших концентраций. Иногда правая часть (увеличение  $RSD$  в области больших концентраций) не наблюдается. Подобные графики очень удобны в аналитической практике.

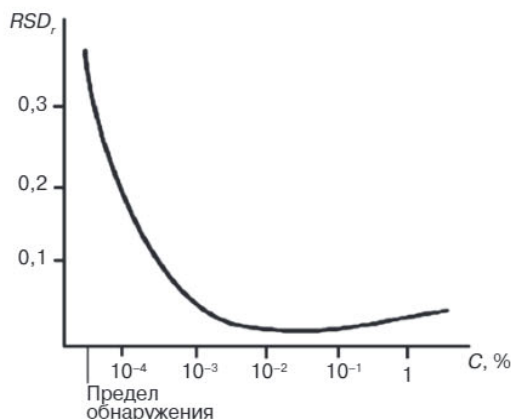


Рис. 3.4. Типичная зависимость стандартного отклонения сходимости  $RSD_r = s_r / \bar{X}$  от концентрации аналита  $C$  для метода анализа с пределом обнаружения  $3 \cdot 10^{-5}\%$

Из опыта известно, что обычно отношение стандартного отклонения сходимости к стандартному отклонению воспроизводимости лежит в диапазоне  $(1:1,2) - (1:5)$  и редко превышает 1,5.

Иногда вместо совместной оценки повторяемости (сходимости) и воспроизводимости методики оценивают повторяемость (сходимость) и промежуточную прецизионность проводят в одной лаборатории. Для этого организуют специальный эксперимент, аналогичный описанному выше, со следующими отличиями:

- все измерения выполняют в одной лаборатории;
- выполняют измерения в нескольких (не менее 8) сериях по 2-4 измерения в серии. При этом в каждой серии используют свое сочетание факторов (время, калибровка, лабортант, прибор), чтобы симитировать разные лаборатории.
- обработку данных проводят по тем же формулам простейшего дисперсионного анализа (фактор — серия) и формулам (3.37).

Найденные оценки  $s_r$  и  $s_{lp}$  фиксируют и приписывают методике КХА в качестве нормативов  $\sigma_r$  и  $\sigma_{lp}$ , выполнение которых требуется от лабораторий валидации (по постановке методики в лаборатории, см. гл. 4).

Такой способ плох по двум причинам:

- почти никогда в одной лаборатории не удастся организовать корректный эксперимент с несколькими приборами, несколькими лабортантами и т. д. — их просто нет, и организаторы подобных экспериментов обычно кривят душой;
- вместо  $\sigma_R$  — величины, объективно характеризующей методику, — величина  $\sigma_{lp}$  является заниженной и разной для разных лабораторий. Поэтому оценка промежуточной прецизионности вместо воспроизводимости не рекомендуется. В случаях, когда организовать корректную оценку воспроизводимости методики невозможно (например, методика в реальности реализована только в одной лаборатории), альтернативой является оценка общей неопределенности как суммы неопределенностей этапов ее выполнения (бюджет неопределенностей, см. разд. 3.5.2.3).

При верификации методики для оценки стандартного отклонения промежуточной прецизионности (и стандартного отклонения сходимости) методики анализа при какой-либо концентрации можно использовать *контрольный материал* (см. выше, разд. 3.2), адекватный рутинным пробам, и анализируют его в нескольких аналитических сериях, проводя несколько измерений в каждой серии. Обычно это делают в ходе контроля стабильности (см. разд. 4.5.1).

### 3.4.4. Оценка правильности

Как было указано в п. 3.3.2, показателем правильности методики анализа является значение систематической погрешности  $\Delta_c$ . Для его оценки необходимо иметь точку отсчета — опорное значение, представляющее собой достаточно точную оценку истинного содержания определяемого вещества в анализируемой пробе. Кроме того, о правильности методики можно судить по результатам исследований, выявляющих систематические погрешности отдельных стадий анализа.

В зависимости от природы анализируемого вещества и особенностей изучаемой методики используются разные способы проверки правильности.

*При наличии адекватных образцов сравнения (или стандартных образцов)* проводят несколько ( $I$ ) параллельных измерений (обозначим их результаты как  $C_i$ ) и рассчитывают среднее значение  $\bar{C}$  и стандартное отклонение  $s$  для каждого образца. Далее для каждого образца проверяют гипотезу  $H_0$ : ( $\mu_1 = \mu$ ), где  $\mu_1$  — математическое ожидание результатов измерений концентрации аналита в образце, полученных исследуемым методом,  $\mu$  — опорное значение. Фактически это гипотеза об отсутствии систематической погрешности. Если гипотеза  $H_0$  не отвергается с принятым уровнем доверительной вероятности, то нет оснований полагать, что разница между средними значениями есть следствие не только случайных, но и систематических погрешностей.

Проверку этой гипотезы проводят аналогично тому, как описано в разд. 2.5.3, с тем отличием, что распределение погрешности аттестованного значения  $C_a$  образца сравнения полагают равномерным в доверительном интервале ( $C_a \pm \Delta C$ ), где  $\Delta C$  — погрешность  $C_a$  [67]. В этом случае величина  $t$  имеет  $t$ -распределение с числом степеней свободы  $f = I - 1$ .

$$t = |\bar{C} - C_a| / \sqrt{\frac{s^2}{I} + \frac{(\Delta C)^2}{3}}, \quad (3.38)$$

где  $(\Delta C)^2/3$  — дисперсия равномерно распределенной случайной величины (см. разд. 2.3.1.3).

Если  $t > t(\bar{P}, f)$ , при выбранном уровне доверительной вероятности  $\bar{P}$  (обычно 0,975 для двустороннего критерия), то гипотезу  $H_0$  отвергают, что означает наличие систематической составляющей погрешности анализа данного образца. В этом случае имеет смысл найти оценку систематической погрешности  $\Delta_c = \bar{C} - C_a$ . Проведя такие измерения для образцов с концентрацией, перекрывающей весь интервал определяемых по изучаемой методике концентраций (обычно достаточно



5–8 образцов), получают представление о систематических погрешностях при разных концентрациях<sup>58</sup>.

Для получения оценки постоянной и пропорциональной составляющих систематической погрешности результаты измерений для нескольких образцов обрабатывают в рамках модели регрессионного анализа

$$C_{a,i} = \alpha + \beta C_a + \varepsilon_i, \quad (3.39)$$

где  $C_{a,i}$  — результат  $i$ -го измерения в  $a$ -м образце;  $\varepsilon_i$  — случайная погрешность.

Проверяют гипотезы о равенстве единице тангенса угла наклона прямой ( $\beta$ ) (3.39) и нулю отрезка, отсекаемого на оси ординат ( $\alpha$ ). Если какая-либо из этих гипотез отвергается, говорят о присутствии, соответственно, пропорциональной или постоянной погрешности. Раздельное рассмотрение этих составляющих погрешности имеет смысл постольку, поскольку они вызываются разными причинами. Так, пропорциональная погрешность чаще всего обусловлена некорректной градуировкой, а постоянная — неполным учетом поправки, определяемой в холостом опыте, либо влиянием матрицы. Соответственно, выявление той или иной составляющей погрешности может указать конкретный путь к снижению общей погрешности методики анализа.

Еще одна причина предпочтительности обработки результатов в совокупности в рамках модели (3.39) связана с тем, что в этом случае появляется возможность выявить систематическую погрешность, незаметную при обработке результатов анализа отдельных образцов. Обычно для проверки правильности проводят 2–4 измерения для каждого образца, и даже значительная систематическая погрешность не может быть выявлена на одном из них. Рассмотрение же всех результатов в совокупности в рамках модели (3.39) значительно повышает надежность выводов.

В отсутствие образцов сравнения можно проводить *проверку правильности по отношению к другой, обычно хорошо изученной методике (метод сравнения)*. Для этого одну или несколько рутинных проб с разной концентрацией определяемого вещества исследуют обеими методиками, проводя по несколько параллельных измерений каждым методом. Результаты обрабатывают аналогично тому, как это делается в способе с использованием образца сравнения с тем отличием, что распределение результатов для метода сравнения считают нормальным.

Этот способ проверки правильности хуже предыдущего, поскольку метод сравнения может иметь собственные систематические погрешности, и в случае выявления расхождений между результатами обоих методов неясно, какой из них (или они оба?) был причиной найденной погрешности. Более того, даже если достоверная систематическая погрешность не выявлена, это не гарантирует ее отсутствие — возможно, это связано с односторонними погрешностями обоих методов анализа. Поэтому в качестве метода сравнения стараются выбрать методику, основанную на других, чем исследуемая, физических принципах регистрации аналитического сигнала. Так, результаты фотометрического определения органических веществ

<sup>58</sup> В таком эксперименте также можно найти зависимости компонент прецизионности от концентрации аналита.



сравнивают с результатами хроматографического или масс-спектрометрического анализа.

Тем не менее обычно считают, что отсутствие достоверно выявленных при использовании метода сравнения систематических погрешностей свидетельствует о действительном их отсутствии (с учетом точности опыта).

Еще один способ проверки правильности методики количественного химического анализа в отсутствие адекватных образцов сравнения — это использование метода «введено-найдено»<sup>59</sup> (англ. *recovery*). В простейшем случае берут «типичную» пробу и часть ее анализируют (обычно несколько раз) по исследуемой методике. Обозначим полученную концентрацию как  $C_0$ . Оставшуюся часть пробы делят на несколько частей, точно измеряют их массу и затем к каждой из этих частей добавляют точно известное количество определяемого вещества («добавка»), причем обычно соблюдают условие: масса первой добавки  $\Delta_1 \approx C_0$ , масса второй  $\Delta_2 \approx 2C_0$  и т. д. После тщательного перемешивания определяют концентрацию в каждой из получившихся смесей, проводя для каждой несколько параллельных измерений. В случае отсутствия систематической погрешности измеренные в пробах с добавками концентрации должны составлять (с учетом погрешности измерений)  $C_1 = C_0 + \Delta_1$ ,  $C_2 = C_0 + \Delta_2$  ... Обозначим реально измеренные средние значения концентраций для проб с добавками как  $\bar{C}_{1,изм}$ ,  $\bar{C}_{2,изм}$  и т. д. Разности  $(\bar{C}_{1,изм} - C_1)$ ,  $(\bar{C}_{2,изм} - C_2)$ , ... могут служить показателем наличия систематической погрешности. Проверяют гипотезу о равенстве найденных экспериментально и рассчитанных концентраций. Если для какой-либо из проб с добавкой эта гипотеза отвергается, то значит, что систематическая погрешность для нее не обнаружена. Также можно рассмотреть все указанные разности концентраций в рамках одной регрессионной модели аналогично тому, как это было сделано в предыдущих случаях. При проверке гипотез необходимо учитывать, что дисперсия погрешности разностей концентраций есть сумма дисперсий погрешностей измерения концентрации в основной пробе и в пробе с добавкой.

*Часто вводят только одну добавку. При этом получают менее надежные результаты.*

Существует также упрощенный графический способ представления результатов проверки правильности методом «введено-найдено». График строят в координатах «добавленные концентрации—соответствующие значения аналитического сигнала» (рис. 3.5). Экстраполируют полученную прямую до пересечения с осью  $X$ . В отсутствие систематической погрешности отрезок от нуля до точки пересечения близок к  $C_0$ ; значительное отличие, как и нелинейность графика, указывает на наличие систематической погрешности.

Заметим, что аналогичный метод используют и при нахождении концентрации аналита в рутинных пробах. В каждую пробу вводят одну или несколько добавок, а концентрацию находят либо расчетным, либо графическим методом. В этом случае метод называется «метод добавок» или «метод стандартных добавок» (*standard addition method*).

<sup>59</sup> В [59] этот способ не вполне корректно назван «метод добавок».

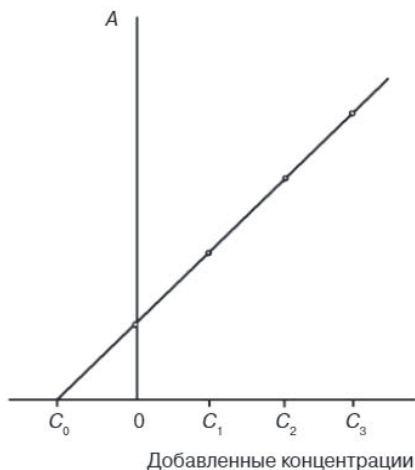


Рис. 3.5. Графическая интерпретация метода добавок (зависимость аналитического сигнала  $A$  от добавленных концентраций  $C_1, C_2$  и т. д.)

К сожалению, методы, связанные с введением добавок, имеют ряд ограничений. Для их использования необходимо иметь возможности вводить добавку в той же форме, в которой определяемое вещество содержится в пробе, и равномерно распределять ее по объему пробы. Гарантировать выполнение этих условий для твердых проб за редким исключением невозможно, поэтому добавки обычно вводят в жидкие и, изредка, в газообразные пробы.

Известны и другие способы проверки правильности методики анализа; разведение проб, введение добавок с предварительным разбавлением пробы [67], с использованием межлабораторного эксперимента [68] и т. д. Однако они малоэффективны и используются сравнительно редко.

Упомянем еще один способ доказательства правильности методики, используемый, в частности, в лаборатории автора. Он состоит в доказательстве правильности выполнения основных этапов анализа.

Методика атомно-абсорбционного определения ванадия в нефтепродуктах состоит из двух основных этапов: переводе образца в раствор (озоление образца в муфельной печи при  $550^\circ\text{C}$  с последующим растворением в смеси минеральных кислот) и измерении концентрации ванадия в полученном растворе на атомно-абсорбционном спектрофотометре. В отсутствие стандартных образцов и образцовой методики для доказательства правильности анализа в целом можно:

- выполнить первый этап двумя способами: предусмотренным методикой и с использованием бомбы (когда исследуемый образец вместе со смесью кислот помещается в бомбу и выдерживается при температуре  $\sim 150^\circ\text{C}$  и давлении 80–100 атм. несколько часов) и сравнить результаты атомно-абсорбционного определения ванадия в растворах, полученных обоими методами. Если результаты в обоих случаях отличаются недостаточно, можно говорить о правильности первого этапа — переводе образца в раствор;
- провести определение ванадия в растворе, полученном основным способом (с озолением образца в муфельной печи) как атомно-абсорбционным, так и независимым методом (например, с помощью масс-спектрометрии с индукционно-связанной плазмой). Совпадение результатов (в пределах погрешности измерений) свидетельствует о правильности второго этапа — определения концентрации ванадия в растворе.

Если оба эксперимента показали совпадение результатов, можно с уверенностью говорить о правильности методики в целом.

Заметим, что проблема правильности анализа намного сложнее, чем описано выше, особенно при анализе проб сложного состава (почвы, геологические образцы, пробы биологического происхождения и т. д.). Дело в том, что из-за наличия скрытых систематических погрешностей, обусловленных влиянием матрицы на результаты анализа, погрешность, оцененная вышеописанными методами, может оказаться разной для образцов (проб) разного состава. В таких случаях правильность характеризуют в ходе исследования образцов (проб), состав которых близок к среднему составу рутинных проб, для анализа которых предназначена методика. Для полной характеристики методики анализа необходимо представлять себе, какие погрешности могут возникнуть из-за влияния состава и свойств индивидуальных проб на результаты анализа.

### 3.4.5. Робастность, селективность и специфичность

В практике анализа эти понятия часто используются для качественной характеристики методики (например, «высокоселективная методика»). Для полной количественной характеристики методики анализа нужно знать зависимость результатов определения концентрации от состава анализируемых проб (часто говорят о влиянии матрицы пробы, о влиянии «третьих» элементов при элементном анализе и т. д.), а также от физико-химических свойств проб и от возможных изменений условий анализа.

Взаимное влияние концентрации определяемых веществ на результаты анализа может быть описано [66] в форме матрицы

$$\begin{pmatrix} C_{1,\text{изм}} \\ C_{2,\text{изм}} \\ \dots \\ C_{N,\text{изм}} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1N} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2N} \\ \dots & \dots & \dots & \dots \\ \gamma_{N1} & \gamma_{N2} & \dots & \gamma_{NN} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_N \end{pmatrix}, \quad (3.40)$$

где  $C_{1,\text{изм}}, C_{2,\text{изм}}, \dots, C_{N,\text{изм}}$  — концентрации определяемых веществ, измеряемые по исследуемой методике;  $C_1, C_2, \dots, C_N$  — истинные концентрации определяемых веществ; коэффициенты прямоугольной матрицы характеризуют взаимное влияние концентраций определяемых веществ на результаты их измерения. Аналогичным образом может быть описано влияние физико-химических свойств пробы и условий проведения анализа на результаты измерения концентраций.

Нахождение коэффициентов в (3.40) весьма трудоемко, но это делается чаще всего при автоматизированных измерениях. Так, в рентгеноспектральном анализе коэффициенты взаимного влияния элементов оценивают путем постановки специального эксперимента (многомерная градуировка) и затем применяют для уточнения результатов анализа. С целью минимизации затрат при изучении влияния состава и свойств проб, а также условий проведения анализов на результаты широко используют планирование эксперимента [34, 69, 70]. Подробнее о методах оценивания селективности методик количественного химического анализа см. [71].



Все описанные выше характеристики в идеале должны были бы определяться для каждой методики количественного химического анализа. Однако на практике подобные исследования чрезвычайно редки из-за их высокой стоимости. В большинстве случаев ограничиваются оценками сходимости, правильности (соответствующие исследования проводятся на имеющихся образцах сравнения типичного состава), а также вида градуировочной характеристики. *При аттестации новых методик КХА оценки воспроизводимости обычно отсутствуют из-за невозможности провести межлабораторный эксперимент. Эти оценки появляются при стандартизации методики, выполняемой после того, как она получила широкое распространение и реализована во многих лабораториях* (см. также выше, разд. 3.4.4).

По результатам исследований методики можно оценить общую допустимую погрешность методики КХА  $\Delta_d$  способами, аналогичными описанным в разд. 3.5.2, если она не задается из других соображений.

На основе этих и других сведений осуществляется собственно аттестация (валидация) методики количественного химического анализа.

### 3.4.6. Аттестация (валидация) методик

Как указано выше, под аттестацией (валидацией) методик мы понимаем процедуру утверждения (узаконения) методик, которая может включать в себя исследование ее метрологических характеристик.

В России аттестации в обязательном порядке подлежат методики измерений (количественного химического анализа), относящиеся к сфере государственного регулирования [2]. Методики, используемые в других областях, например в металлургии, аттестуются в добровольном порядке. Тем не менее это делается достаточно часто — результаты анализа, выполненного по аттестованной методике, почти всегда вызывают больше доверия.

В России аттестацию методик КХА осуществляют организации метрологического профиля и другие структуры, разрабатывающие и применяющие методики химического анализа. При этом организация, аттестующая методику, относящуюся к сфере государственного регулирования, должна иметь государственную аккредитацию.

В других странах подходы могут отличаться от принятого в России. В ряде стран аттестацией методик занимаются различные организации, в том числе и общественные. Зачастую менее формализована и система аккредитации этих организаций. Иногда она вообще отсутствует, и многое зависит от их авторитета. В качестве примера приведем ASTM International (American Society for Testing and Materials) — некоммерческую организацию, разрабатывающую и издающую стандарты для материалов, продуктов, систем и услуг, в том числе и методики КХА. Методики ASTM уважаемы и широко используются во всем мире.

Рассмотрим процедуру аттестации, принятую в Российской Федерации.

На аттестацию представляются: исходные требования на разработку методики; документ (проект документ), регламентирующий методику; результаты исследования методики (если оно проводилось).

Организация, осуществляющая аттестацию, проводит метрологическую экспертизу представленных документов как по форме, так и по содержанию. Прове-



ряется корректность проведенных исследований, способов расчетов, соответствие метрологических характеристик методики ее назначению и т. д. При положительном результате аттестации документ, регламентирующий методику, утверждается, а сведения о ней заносятся в соответствующие реестры. После этого методика считается аттестованной и допускается к применению.

При всей привлекательности описанной процедуры аттестации требование *обязательности* аттестации (валидации) методик имеет и отрицательные последствия. Прежде всего, аттестация методики влечет за собой дополнительные затраты. Требование обязательной аттестации зачастую сдерживает технический прогресс: вновь разработанные методики, превосходящие аттестованные по метрологическим характеристикам или по экономичности, не могут использоваться (по крайней мере легально) до их аттестации. Иногда такое положение продолжается годами. Кроме того, наличие аттестованной методики часто создает иллюзию, что с ее применением лаборатория решает все проблемы обеспечения качества анализа. Однако такое впечатление обманчиво: в отличие от «классических» измерений методика при реализации ее в конкретной лаборатории может не дать ожидаемых метрологических характеристик или они могут меняться во времени непредсказуемым образом.

Первые две причины привели к отсутствию де-факто аттестации методик, предназначенных для клиничко-диагностических исследований — динамично развивающейся области, в которой едва ли не ежедневно появляются новые методы и методики КХА и других исследований, а конкурентная борьба весьма остра. В то же время «естественный отбор» методик происходит в ходе внешней оценки качества (см. разд. 4.4), получившей в лабораторной медицине широкое распространение.

Проблемы нестабильности методик анализа и трудности их реализации в конкретной лаборатории с метрологической точки зрения в значительной степени решаются с помощью контроля качества результатов анализа «по конечному результату». Важнейшие из таких методов — внутрилабораторный контроль качества (разд. 4.5) и внешняя оценка качества химического анализа (разд. 4.6) — описаны в гл. 4.

### 3.5. Прослеживаемость и неопределенность результатов анализа

Важнейшими понятиями, характеризующими результаты измерений вообще и КХА в частности, являются прослеживаемость и неопределенность. Хотя эти термины появились у нас недавно, они уже получили широкое распространение из-за важности связанных с ними понятий. Ниже рассмотрены прослеживаемость и неопределенность результатов КХА и взаимосвязь этих понятий.

#### 3.5.1. Прослеживаемость результатов измерений в КХА

Рассмотрим понятие «прослеживаемость», опираясь на [65, 72].

*Прослеживаемость, или метрологическая прослеживаемость (англ. «Traceability») есть свойство измерений, посредством которого результат измерения может быть соотнесен с основой для сравнения через документированную непрерывную цепь кали-*

*бровок, каждая из которых вносит вклад в неопределенность измерений.* Речь идет о реализации второй аксиомы метрологии, см. разд. 1.1). Для того чтобы что-то измерить, необходимо сравнить измеряемую величину с другой, принятой за точку отсчета. Заметим, что принципиальной новизны в этом нет — лишь новый термин и более четкие формулировки. Фактически прослеживаемость (хотя раньше ее так и не называли) лежит, в частности, в основе системы обеспечения единства измерений путем передачи размера физических величин от первичного эталона к образцовым и рабочим средствам измерений (см. разд. 1.4), в которой «калибровка» выполняется в ходе поверок, а «основа для сравнения» — один из эталонов высокого уровня<sup>60</sup>. Однако прослеживаемость в КХА имеет свою специфику, определяемую особенностями метрологии химического анализа (см. разд. 3.1.2). Рассмотрим, что может быть «основой для сравнения» и как выполняется «калибровка» в КХА.

Выделяют три вида «опорных основ» (реперов), к которым может проследиваться результат измерений:

- единица измерения, воспроизводимая по ее определению;
- эталон, например, аттестованный стандартный образец;
- [образцовая] методика измерения, которая полностью определена и согласована на международном уровне.

Как было показано выше, в КХА основой для сравнения может быть достаточно чистое вещество, взятие навески которого позволяет воспроизвести размер моля (по крайней мере в случаях, когда известна формула вещества). Другими словами, навеска чистого вещества с известной формулой воспроизводит единицу измерения (моль) по ее определению. Практически все методики КХА представляют собой косвенное измерение (см. разд. 1.3.1), в котором размер искомой величины находят по формуле либо в ходе градуировки — в обоих случаях размер передается от навески чистого вещества. Казалось бы, больше ничего не нужно. Однако это не так — особенности метрологии химического анализа, особенно наличие скрытых систематических погрешностей (см. разд. 3.1.2), затрудняющее оценку погрешности (неопределенности) методики, вынуждают к использованию других путей прослеживаемости.

То есть при химическом анализе в большинстве случаев воспроизведения единицы измерения недостаточно, и требуются другие точки отсчета, играющие роль эталона, такие как стандартный образец [сложного состава] или методика измерения либо их комбинации.

Конечно, результат измерения должен быть всегда прослежен до соответствующего стандартного образца, образцовой методики и т. д., и лаборатория должна быть способна продемонстрировать это. Лаборатория может вывести свою собственную цепь прослеживаемости, изучая документацию для своих рутинных методик, оборудования и калибраторов. Однако в реальности осуществление и демонстрация метрологической прослеживаемости в КХА часто оказывается непростой задачей. Это связано с особенностями метрологии химического анализа, приводящими к использованию нескольких способов получения результата измерения для одной

<sup>60</sup> Понятие «единство измерений», при всей близости к понятию «прослеживаемость» шире — оно включает также использование узаконенных единиц измерений.



и той же измеряемой величины. Сложный состав исследуемого материала означает, что часто требуется непростая предварительная обработка и очистка пробы, что делает затруднительными прямые сравнения между эталонами и исследуемыми пробами. Особые проблемы создает наличие скрытых систематических погрешностей. В результате, хотя в основном результаты КХА прослеживаются к чистым веществам и стандартным образцам состава, последние во многих случаях, в свою очередь, аттестуются методом межлабораторного эксперимента.

**Пример.** При определении кальция в биологических жидкостях (сыворотке крови, моче) в обычных лабораториях часто используются спектральные методы (атомно-абсорбционный или атомно-эмиссионный анализ). При этом проводится сложная подготовка пробы (озоление пробы и растворение ее в минеральных кислотах), а при измерении выполняется градуировка по водным растворам чистых солей кальция. Для проверки правильности результатов анализа используют стандартные образцы состава, опорное значение для которых получается либо методом межлабораторного эксперимента, либо с помощью образцовой методики — методом масс-спектрометрии с изотопным разбавлением. Понятно, что в данном случае результаты измерений могут прослеживаться разными путями к разным «точкам отсчета», и оптимальный выбор того или иного пути не очевиден.

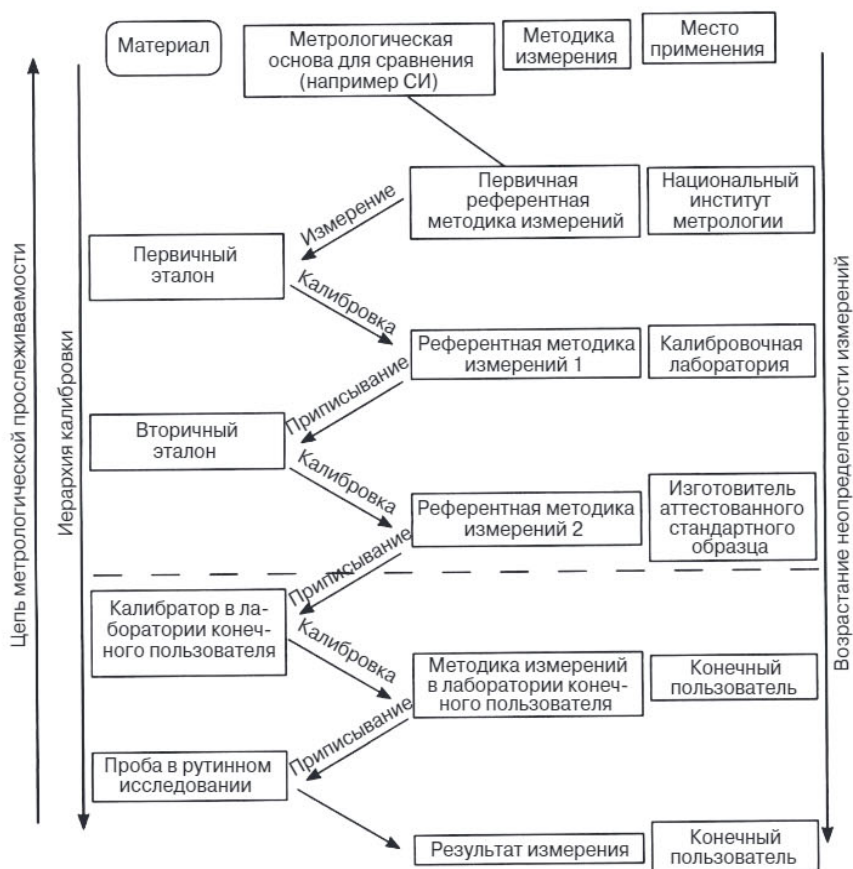


Рис. 3.6. Пример цепи прослеживаемости, включающей референтные (образцовые) методики измерений. Подобные схемы редко реализуются на практике

Часто результаты КХА прослеживаются в конечном счете к некоторой методике, лишь в той или иной степени «привязанной» к чистым веществам через процедуру градуировки.

Прослеживаемость в химическом анализе может осуществляться как по простым схемам (непосредственно к чистому веществу в ходе градуировки, по схеме, показанной на рис. 3.1, так и более сложными путями (см., например, рис. 3.6). Оптимальным является путь, приводящий к наименьшей погрешности (неопределенности) результатов, однако из вышесказанного ясно, что точно оценить ее непросто и зачастую невозможно. Выбор, часто неформальный, остается за лабораторией.

Изучение цепочки прослеживаемости и возникающих при этом погрешностей может быть необходимо, в частности, при оценке неопределенности результатов измерений.

### 3.5.2. Оценка неопределенности измерений и представление их результатов заказчику

В гл. 1 было рассмотрено понятие «погрешность измерений». Наряду с ним в последние годы используется понятие «неопределенность». Неопределенность и связанные с ней величины (стандартная неопределенность, расширенная неопределенность и т. д.), в последнее время все шире используются при представлении результатов измерений, особенно в европейских странах. То же происходит и в России. Так, понятие неопределенности фигурирует в ряде нормативных документов (например, в [73–75]), причем основополагающий стандарт [73] прямо требует от лабораторий иметь процедуры оценки неопределенности результатов измерений.

#### 3.5.2.1. Что такое неопределенность

Понятие «неопределенность» (англ. «Uncertainty») в лабораторном деле появилось более 30 лет назад.

Фундаментальным понятием классической теории измерений является *погрешность*  $\delta_i = X_i - \mu$  – отклонение результата измерения  $X_i$  от истинного значения измеряемой величины  $\mu$  (см. разд. 1.3.5). Погрешность возникает из-за несовершенства процесса измерений. Хотя погрешность не может быть точно известна (из-за неизвестности истинного значения), это понятие удобно использовать для статистического описания процесса измерений. Распределение погрешности совпадает с распределением результатов измерений  $X$  с точностью до начала координат (см. рис. 3.7).

Рассмотрим теперь, как определяется неопределенность.

Согласно [76, 77], это «*Параметр, связанный с результатом измерения и характеризующий разброс значений, которые с достаточным основанием могут быть приписаны измеряемой величине. ... Этим параметром может быть, например, стандартное отклонение (или кратное ему число) или ширина доверительного интервала*».

Близкое по смыслу определение дано в [1]: «*Неопределенность измерений есть неотрицательный параметр, характеризующий рассеяние значений величины, приписываемых измеряемой величине на основании используемой информации*», причем



оговаривается, что 1) «неопределенность включает составляющие, обусловленные систематическими эффектами...» и 2) «параметром может быть, например, стандартное отклонение, называемое стандартной неопределенностью измерений (или кратное ему число) или половина ширины интервала с установленной вероятностью охвата».

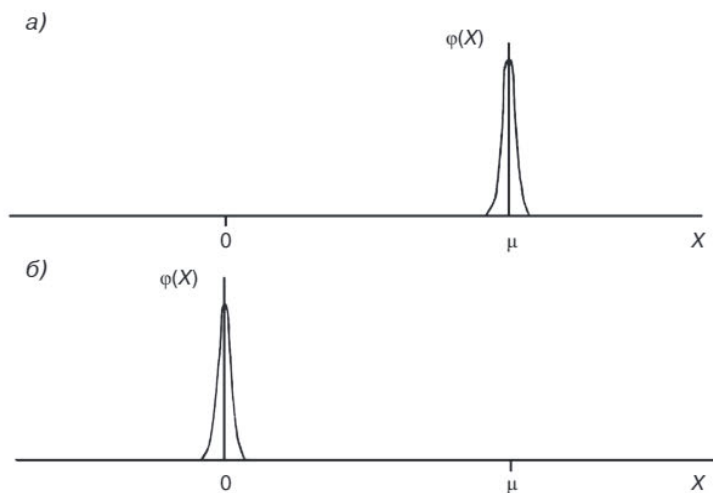


Рис. 3.7. Плотность распределения измеряемой величины  $X$  (а) и погрешности измерений  $\delta = X - \mu$  (б).  $\phi(X)$  — плотность распределения

Сравним понятия «погрешность» и «неопределенность», следуя [78].

Неопределенность есть параметр распределения центрированной случайной величины, представляющей собой разность между истинным значением измеряемой величины и результатом измерений, то есть величины, совпадающей по модулю с погрешностью измерений, но противоположной ей по знаку, то есть это параметр распределения величины  $(\mu - X)$ . Плотность такого распределения зеркальна по отношению к плотности распределения погрешности измерений (рис. 3.8а) и, следовательно, они совпадают друг с другом при симметричных функциях плотности распределения (рис. 3.8б), с которыми мы почти всегда имеем дело. Рассматривая литературные данные, можно сказать, что *во всех случаях* в связи с неопределенностью рассматриваются *симметричные плотности распределения результатов измерений* (результаты измерений рассматриваются как выборка из нормально распределенной генеральной совокупности), а в качестве параметра, упомянутого в определении неопределенности, *всегда* рассматриваются *стандартные отклонения*. Таким образом, все данные выше определения неопределенности оказываются эквивалентными, а между описанием результатов измерений с использованием погрешности (представляемой как параметр распределения) и с использованием неопределенности имеется полное соответствие.

В табл. 3.2 приведены основные термины, используемые при описании результатов измерений с помощью неопределенности, и аналогичные величины при использовании погрешности.

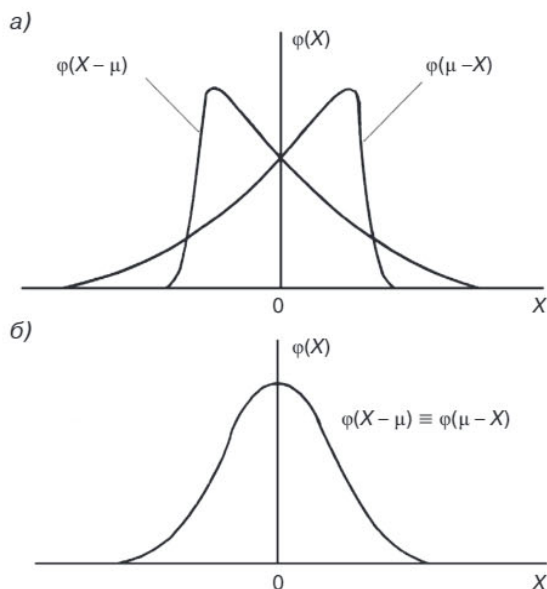


Рис. 3.8. Плотность распределения погрешности результата измерений  $\delta = (X - \mu)$  и величины  $(\mu - X)$ : несимметричное распределение измеряемой величины (а), симметричное (б)

Таблица 3.2. Неопределенность и погрешность

Описание результатов измерений с использованием:		Описание характеристики
НЕОПРЕДЕЛЕННОСТИ	ПОГРЕШНОСТИ	
Среднее значение или медиана	Среднее значение или медиана	Результат измерения (точная оценка)
Стандартная неопределенность $u(X)$	Стандартное отклонение результатов измерений	Разброс, характеризующий результаты измерений
Суммарная стандартная неопределенность $u_c(X)$	Стандартное отклонение полной (суммарной) погрешности	Разброс, характеризующий все виды ошибок вместе
Расширенная неопределенность $U = k u_c(X)$ , $k$ — коэффициент охвата (см. ниже)	Полуширина доверительного интервала, рассчитанного с учетом стандартного отклонения полной (суммарной) погрешности	Полуширина интервала, в котором лежит измеряемая величина при единичном измерении

Приведенные в табл. 3.2 величины рассчитываются по-разному в разных случаях. В табл. 3.3 даны формулы для расчета погрешности и неопределенности для случая, когда они определяются только случайным разбросом измерений и погрешностью (неопределенностью) стандартного образца, к которому прослеживаются их результаты (предполагается, что результаты измерений  $X_1, X_2, \dots, X_i, \dots, X_N$  распределены нормально). Следует иметь в виду, что при представлении стандартной либо расширенной неопределенности часто говорят просто «неопределенность».

Таблица 3.3. Сравнение расчетных формул

Расчетные формулы, применяемые при описании результатов измерений использованием:	
НЕОПРЕДЕЛЕННОСТИ	ПОГРЕШНОСТИ
Среднее значение: $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$	Среднее значение: $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$
Стандартная неопределенность: $u(X) = +\sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1}}$	Стандартное отклонение результатов единичных измерений: $s_r = +\sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1}}$
Суммарная стандартная неопределенность: $u_c(X) = +\sqrt{u^2(X) + u^2(\hat{\delta}) + \sum_{j=1}^J C_j^2 u^2(z_j)}$	Стандартное отклонение полной (суммарной) погрешности: $s_{\text{сумм}} = +\sqrt{s_r^2 + \frac{(\Delta C)^2}{3}}$
Расширенная неопределенность: $U = k u_c(X),$ где $k$ — коэффициент охвата	Полуширина доверительного интервала для стандартного отклонения полной (суммарной) погрешности: $k s_{\text{сумм}}$ , где $k$ — величина, по смыслу аналогичная коэффициенту охвата
Интервал, в котором лежит измеряемая величина: $\{\bar{X} \pm U\} = \{\bar{X} \pm k u_c(X)\}$	Доверительный интервал: $\{\bar{X} \pm k s_{\text{сумм}}\}$

Использованы следующие стандартные обозначения:

- $\Delta C$  — полуширина интервала, в котором лежит принятое опорное значение для стандартного образца (см. разд. 3.2.4);
- $u^2(z_j)$  и  $C_j^2$  — соответственно, неопределенность, характеризующая распределение  $j$ -го мешающего компонента, и коэффициент, учитывающий его вклад в неопределенность определения аналита;
- величина  $u^2(\hat{\delta})$  учитывает неточность аттестации стандартных образцов и равна  $(\Delta C)^2/3$ ;
- член  $\sum_{j=1}^J C_j^2 u^2(z_j)$ , описывающий вклад неопределенности, обусловленной матричными эффектами, в отечественной литературе обычно не рассматривается (полагают, что эти эффекты малы);
- коэффициент охвата  $k = u(\bar{P})$ , где  $\bar{P}$  — доверительная вероятность, при расчете расширенной неопределенности выбирают равным 2 или 3, что соответствует  $\bar{P} = 0,954$  и  $\bar{P} = 0,997$ . Иногда используют и другие значения  $k$ , например,  $k = 1,96$  соответствует  $\bar{P} = 0,95$ . Если стандартная неопределенность оценивалась по небольшому числу измерений ( $N < 8$ ),  $k$  выбирают равным процентной точке распределения Стьюдента (см. Приложение 3) для  $\bar{P} = 0,95$ .

Как видно из табл. 3.3, расчетные формулы для соответствующих друг другу величин, используемых при описании результатов измерений с использованием суммарной погрешности и неопределенности, практически совпадают.

### 3.5.2.2. Причины введения неопределенности и необходимость ее использования

Основные причины, по которым вводится понятие «неопределенность», следующие.

1. При использовании понятия «погрешность» отсчет доверительного интервала, по определению, должен был бы вестись от математического ожидания



$\mu$ <sup>61</sup> (см. разд. 2.1.2), а величину  $\bar{X}$  берут не от хорошей жизни — истинное значение неизвестно, а  $\bar{X}$  есть наилучшая оценка для  $\mu$  [79]. Применять понятие «неопределенность» при описании результатов измерений более естественно — в определениях всех параметров фигурируют только наблюдаемые величины. Запись интервала в форме (*среднее значение  $\pm$  расширенная неопределенность*) выбирается лишь из соображений удобства.

2. Способы оценки интервала, в котором (с принятой доверительной вероятностью) лежит истинная величина, весьма разнообразны и детально прописаны в [77] и других документах, использующих понятие «неопределенность»<sup>62</sup>. В частности, учитываются реально имеющие место, но зачастую игнорируемые в отечественных нормативных документах скрытые, или не выявленные, систематические погрешности (см. разд. 3.2). Возможно даже использование экспертных оценок.

3. Использование неопределенности позволяет наглядно решать вопрос о соответствии (несоответствии) измеренной величины установленным нормам. Если значение нормы не перекрывается расширенной неопределенностью результата измерения, то, основываясь на этом результате можно делать надежное заключение о соответствии (несоответствии) объекта испытания этой норме (см. разд. 3.5.2.4.2). Правда, то же относится к корректно рассчитанному доверительному интервалу при использовании погрешности.

4. Понятие «погрешность» исходно не являлось на Западе настолько же распространенным, как в России. Поставленные перед необходимостью оценивать интервал, в котором лежит истинное значение, зарубежные специалисты за основу взяли «неопределенность» — у них был выбор.

Необходимость применения понятия «неопределенность» в России несомненна. Ее всегда рассчитывают в лабораториях, работающих в экспортно-ориентированных областях — представление сведений о неопределенности результата является стандартным требованием зарубежных партнеров. Но и для других отечественных лабораторий использование понятия «неопределенность» постепенно становится фактически обязательным. Это справедливо, прежде всего, для аккредитованных лабораторий. В настоящее время все системы аккредитации основываются на требованиях документа [73], в котором установлено, что испытательные лаборатории должны иметь и применять процедуры оценки неопределенности измерений. О том же говорят документы ряда международных организаций ILAC, EURACHEM, CITAC (см. разд. 1.6.5) и т. д. Из термина, которым ранее оперировал узкий круг метрологов, «неопределенность» превратилась в понятие, неотъемлемое от результата любого измерения. Другими словами, используя неопределенность, мы говорим на одном языке с остальным миром.

Неопределенность должна оцениваться во всех случаях, за исключением качественного анализа (да и то только до выхода соответствующего Руководства ILAC) [80]. При этом требование о представлении оцененной неопределенности является обязательным в следующих случаях:

<sup>61</sup> В отсутствие систематических погрешностей истинное значение совпадает с математическим ожиданием генеральной совокупности (см. разд. 2.1.1).

<sup>62</sup> Такие же формулы справедливы и при использовании понятия «погрешность».



- если этого требует Заказчик;
- если из-за отсутствия в протоколах сведений о неопределенности возможна неправильная интерпретация результата;
- когда неопределенность превышает рассчитанную по данным, приведенным в методике. Например, лаборатория выявила для каких-то объектов матричных эффекты, не отмеченные в методике.

В то же время стандарты [73, 80] указывают, когда представление в протоколах испытаний сведений об оцененной неопределенности не является обязательным. Так, лаборатория может не приводить в протоколах испытаний сведения о неопределенности, если заказчик измерений отказался от получения этой информации. Не обязательно представлять сведения о неопределенности в случаях, когда результаты получены по широко известной стандартной методике, содержащей значения оцененной неопределенности для этих результатов, и применение методики конкретной лабораторией не изменяет эту неопределенность. Однако следует иметь в виду, что *приведение в протоколах измерений сведений о неопределенности есть признак хорошего тона*.

### 3.5.2.3. Способы оценки неопределенности результатов КХА

Известно несколько способов оценки неопределенности результатов измерений. Как и метрологические характеристики, оцениваемые при исследовании методик (см. разд. 3.3), *оценки неопределенности имеют смысл только в применении к конкретной методике КХА*.

#### 3.5.2.3.1. Оценка неопределенности по известным метрологическим характеристикам методики

Если методика исследована (то есть для нее в целом исследована правильность и найдены компоненты прецизионности, см. разд. 3.3), то неопределенность можно оценить с использованием алгоритма, представленного на рис. 3.9 [81].



Рис. 3.9. Алгоритм оценки неопределенности, рекомендованный [80]

Вначале выясняют, имеется ли у методики достоверно зафиксированная систематическая погрешность (п. 1 на рис. 3.9). Если методика корректно исследована и аттестована в соответствии с [67], то такой погрешности она не имеет. Если постоянная систематическая погрешность выявлена, например, при валидации методики в лаборатории, ее можно учесть, внося соответствующую поправку в результаты измерений (при этом необходимо будет учесть неопределенность самой этой поправки). Если поправка не вносится, придется учитывать вклад систематической погрешности в суммарную стандартную неопределенность. Как это делается, описано в 3.5.2.3.2.

Далее берут оценки повторяемости (сходимости) и воспроизводимости методики, полученные в ходе ее аттестации (о способах их получения см. разд. 3.4.3). В большинстве случаев эти оценки содержатся в тексте методик КХА (п. 2 на рис. 3.9).

Затем рассматривают методику с точки зрения наличия еще каких-то факторов, которые могут вносить вклад в неопределенность результатов и не учтены в методике КХА (п. 3 на рис. 3.9). Это могут быть, например, не учтенные в методике матричные эффекты (см. Пример 3). Если такие факторы выявлены лабораторией, их вклад в суммарную стандартную неопределенность придется учесть, что можно сделать либо на основе экспертной оценки (как в Примере 3), либо методами, изложенными в разд. 3.5.32.3.2. (п. 4 на рис. 3.9).

Для абсолютного большинства широко используемых методик вкладом систематических погрешностей и неучтенных в методике факторов в погрешность (неопределенность) результатов можно пренебречь и все определяется воспроизводимостью методики. *В таком случае суммарная стандартная неопределенность равна стандартному отклонению воспроизводимости методики  $s_R$ .*

Ниже даны примеры оценки расчета неопределенности и корректного представления результатов измерений, соответствующие рекомендациям [80, 81] и приближенные к реальной практике.

**Пример 1.** Лаборатория определяет pH молока по ГОСТ 26781-85 [82]. В тексте стандарта указано, что «предел возможных значений погрешности измерений =  $\pm 0,04 pH$  для принятой [доверительной] вероятности  $\bar{P} = 0,95$ ». Понятно, что эта величина и есть расширенная неопределенность, и дополнительные оценки и расчеты не нужны.

Пусть измеренное значение  $pH = 6,96$  [единиц  $pH$ ]. Тогда результат может быть записан либо как 6,96, либо как  $6,96 \pm 0,04$  [единиц  $pH$ ].

**Пример 2.** Лаборатория проводит определение содержания меди в поваренной соли методом атомной абсорбции в пламени ацетилен-воздух по ГОСТ 26931-86 «Сырье и продукты пищевые. Методы определения меди» [83]. В методике указано, что «допускаемое расхождение между результатами испытаний, выполненных в двух разных лабораториях, не должны превышать 40% по отношению к среднеарифметическому значению при доверительной вероятности 0,95». Другими словами, известна величина предела воспроизводимости  $\bar{R} = 0,4\bar{X}$ . Поскольку  $R = 2,77 s_R$  (см. разд. 3.3.1), стандартное отклонение воспроизводимости (норматив)  $\sigma_R = R / 2,77 = 0,144\bar{X}$  (можно записать также « $RSD_R = \sigma_R / \bar{X} = 14,4\%$ »).

Известно, что при аттестации данной методики был проведен межлабораторный эксперимент, и нет оснований считать, что при этом не были учтены какие-либо дополнительные источники неопределенности. Поэтому для оценки расширенной неопределенности результатов достаточно использовать только сведения о воспроизводимости.

Тогда суммарная стандартная неопределенность  $uc(X)$  равна стандартному отклонению воспроизводимости  $\sigma_r = 0,144\bar{X}$ , а расширенная неопределенность при коэффициенте охвата  $k = 2$  составляет  $U = ku_c(X) = 2 \cdot 0,144\bar{X}\% = 0,288\bar{X}$  ( $\bar{X}$  – результат измерения в мг/кг).

Предположим, окончательно приводимый результат определения концентрации меди (см. разд. 4.4) равен 102 мг/кг. Тогда  $u_c(X) = 14,7$  мг/кг и  $U = 29,4$  мг/кг  $\approx 29$  мг/кг. Поскольку при представлении результата допускается использование как стандартной неопределенности, так и расширенной неопределенности, записать его можно двумя способами, либо:

«Результат определения меди 102 мг/кг при стандартной неопределенности 29 мг/кг», либо:

«Результат определения меди  $(109 \pm 29)$  мг/кг (при коэффициенте охвата 2)».

Предпочтительным – более наглядным – является второй вариант.

**Пример 3.** Лаборатория проводит определение объемной доли бензола в бензине для широкого круга заказчиков по методике ГОСТ Р 51930 [84]. В тексте методики отсутствуют сведения о неопределенности результатов, но имеются следующие данные, по которым можно провести ее оценку:

- окончательно приводимый результат есть результат единичного измерения;
- диапазон определения объемной доли бензола от 0,1 до 5,0%;
- предел повторяемости  $r = 0,08\%$  (абсолютных);
- предел воспроизводимости  $R = 0,18\%$  (абсолютных);
- «тяжелые ароматические соединения» оказывают мешающие воздействия на результат определения бензола;
- требуются специальные меры для обеспечения стабильности образца в процессе хранения и анализа;
- правильность результатов в межлабораторном эксперименте не оценивалась.

Как указано, сведения о пределах повторяемости и воспроизводимости приведены в абсолютных величинах. Для вычисления неопределенности перейдем от пределов повторяемости и воспроизводимости к соответствующим стандартным отклонениям (см. разд. 3.3.1). Получаем  $\sigma_r = r/2,77 = 0,029\%$  и  $\sigma_R = R/2,77 = 0,065\%$ .

Если бы воспроизводимость охватывала бы все возможные источники неопределенности (погрешности), то ее значение являлось бы собственно неопределенностью и можно было бы сразу переходить к вычислению расширенной неопределенности  $U$ , как это было сделано в Примере 2. Тогда мы получили бы следующие величины в крайних точках диапазона определяемых величин (при коэффициенте охвата  $k = 2$ ):

Объемная доля бензола (%)	$u_c(X) = \sigma_r$	$U = 2u_c(X)$	Запись результата (в %)
0,1	0,065	0,13	$0,10 \pm 0,13$
5,0	0,065	0,13	$5,00 \pm 0,13$

Интересно, что в методике содержится противоречие – концентрацию бензола на нижней границе определяемых концентраций достоверно измерить невозможно. Заметим также, что если бы методика предусматривала два измерения для каждой рутинной пробы ( $N = 2$ ), неопределенность была бы практически такой же:

Объемная доля бензола (%)	$u = \sqrt{\sigma_r^2 - \sigma_r^2(1 - 1/N)}$	$U = 2u_c(X)$	Запись результата (в %)
0,1	0,062	0,12	$0,10 \pm 0,12$
5,0	0,062	0,12	$5,00 \pm 0,12$



Это иллюстрирует тот факт, что почти всегда увеличение числа измерений для рутинной пробы не приводит к сколько-нибудь заметному снижению неопределенности, а если выполнят несколько измерений — то для того, чтобы иметь возможность устранять влияние грубых промахов на окончательно приводимый результат (путем проверки приемлемости, разд. 4.4).

Однако в данном случае расчет неопределенности на этом не заканчивается. В методике отмечено влияние нестабильности образцов и матрицы проб (присутствие тяжелых ароматических соединений) на результаты анализа. Влияние матрицы предлагается учитывать, анализируя пробу холостого опыта — идентичную анализируемой пробе, но без бензола. Однако это невозможно, если лаборатория исследует пробы различного (и, как правило, неизвестного) состава. При аттестации методики влияние вариации матрицы пробы на результаты измерений не учитывается — используются однородные контрольные материалы одного состава, и величина  $s_R$  не учитывает этот фактор.

К сожалению, количественные данные по влиянию нестабильности проб и матричных эффектов на результаты анализа (и на их неопределенность) отсутствуют. Можно попытаться оценить влияние этих факторов на неопределенность с использованием экспертных оценок («профессиональных суждений», [77]). Такие оценки дают  $u_{\text{матр}} = 0,1X$  для неопределенности, определяемой матричными эффектами, и  $u_{\text{нест}} = 0,02X$  для неопределенности, связанной с нестабильностью проб ( $X$  — результат измерения).

Составляющие неопределенности выражены стандартными отклонениями. Поэтому для расчета суммарной неопределенности при нескольких факторах, определяющих неопределенность, используются правила сложения погрешностей (см. разд. 2.6.1). В данном случае имеем три независимых фактора ( $u_R = s_R$ ,  $u_{\text{матр}}$  и  $u_{\text{нест}}$ ), и суммарная неопределенность равна

$$u_c(X) = \sqrt{u_R^2 + u_{\text{матр}}^2 + u_{\text{нест}}^2}. \quad (3.41)$$

Получаем следующие результаты на концах диапазона определяемых концентраций.

Объемная доля бензола (%)	$u_R$	$u_{\text{матр}}$	$u_{\text{нест}}$	$u_c(X)$	$U = 2u_c(X)$	Запись результата (в %)
0,1	0,065	0,01	0,002	0,066	0,132	$0,10 \pm 0,13$
5,0	0,065	0,5	0,1	0,595	1,190	$5,0 \pm 1,2$

Как и следовало ожидать исходя из принципа пренебрежения малыми погрешностями (разд. 2.6.1), вкладом нестабильности можно было пренебречь.

При низких содержаниях бензола дополнительные (неучтенные в процессе аттестации методики) факторы практически не влияют на неопределенность результата, а при высоких содержаниях их вклад в неопределенность результата резко возрастает. Таким образом, игнорирование дополнительных факторов может привести к переоценке качества результатов, и, как следствие, к ошибочному заключению о качестве бензина.

Описанный способ основан на том, что корректно полученные оценки правильности и компонент прецизионности методики КХА в целом включают большинство составляющих неопределенности. Если такие оценки (чаще всего полученные в ходе ее аттестации) отсутствуют, приходится использовать другие способы.

### 3.5.2.3.2. Оценка суммарной неопределенности как суммы неопределенностей этапов выполнения методики КХА

Такой подход подробно рассмотрен, в частности, в руководстве [77]. Для его применения необходимо теоретически или экспериментально оценить все составляющие



неопределенности и просуммировать их по правилам, описанным в разд. 2.6.2. Заметим, что этот способ — едва ли не единственно возможный при оценке неопределенности эталонов высших разрядов (см. разд. 1.4).

Для оценки неопределенности по данному способу необходимо прежде всего детально проанализировать методику измерений и найти все возможные источники неопределенности. В метрологии это называют построением *модели измерений*. Для КХА такой анализ обычно делается последовательно: выделяют этапы выполнения анализа (отбор проб, подготовка пробы, измерение аналитического сигнала и т. д.), и для каждого этапа определяют составляющие неопределенности. Именно отсутствие *полной* модели измерений наряду с трудностями оценки составляющих неопределенности затрудняет нахождение суммарной неопределенности как суммы неопределенностей всех этапов выполнения методики. Другими словами, поскольку мы очень часто не знаем всех основных источников неопределенности, достаточно точно рассчитать таким способом общую неопределенность для таких сложных измерений, как химический анализ, не всегда удастся.

Перечень источников неопределенности с их значениями часто называют «бюджет неопределенности». После оценки составляющих неопределенности (фактически это стандартные отклонения) их суммируют по правилам сложения погрешностей (см. разд. 2.6.1).

Модель измерений представляют в разных формах: словесное описание, диаграммы «причина-следствие» (см. ниже) и т. д. Далее необходимо все выделенные составляющие неопределенности представить количественно в унифицированной форме — как стандартные отклонения.

При этом пользуются следующим набором правил.

1. Если неопределенность, вносимую каким-либо фактором, оценивали, выполняя многократные измерения, стандартное отклонение результатов этих измерений и есть соответствующая стандартная неопределенность. Рекомендуется выполнять не менее 8 измерений (хотя, конечно, лучше больше, 10–15 или более). Если же число измерений невелико ( $N < 8$ ), результат необходимо разделить на  $M_v$  — момент отношения  $s/\sigma$  (см. разд. 2.2.2 и табл. 2.1). Как правило, для последующих вычислений рассчитывают как стандартное отклонение  $s$ , так и относительное стандартное отклонение  $RSD = s / \bar{X}$ .

**Пример.** Для оценки вклада неопределенности титрования одним лаборантом в одних условиях было выполнено  $N = 7$  измерений. Рассчитанное стандартное отклонение составило 0,0242 мл при среднем значении 5,224 мл. Разделив 0,0242 на 0,959 (см. табл. 2.1, получаем несмещенную оценку  $s = 0,0252$  и  $RSD = 0,0252/5,224 = 0,00482$ . Это и есть стандартная неопределенность.

2. Если известен интервал, в пределах которого лежит измеряемая величина, и никакой дополнительной информации нет, случайную величину полагают *равномерно распределенной* в этом интервале (см. разд. 2.3.1.3). Такова, например, ситуация с мерами и со стандартными образцами.

**Пример.** Необходимо оценить неопределенность, вносимую при проверке правильности методики стандартным образцом (СО) водного раствора катиона свинца. В паспорте СО указано, что концентрация свинца (опорное значение) составляет 10,0 г/дм<sup>3</sup>, а относительная погрешность равна 1% при  $\bar{P} = 0,95$ . Другими словами, известно, что концентрация катиона свинца лежит в интервале  $(10,0 \pm 0,1)$  г/дм<sup>3</sup>, и другой информации нет, поэтому

принимая распределение концентрации в СО равномерным в этом интервале. Соответственно, дисперсия равна  $s^2 = \Delta^2/3 = (0,1)^2/3 = 0,00333 \text{ (г/дм}^3)^2$  (см. разд. 2.3.1.3), стандартное отклонение  $s = \sqrt{0,00333} \approx 0,06 \text{ г/дм}^3$ . Это и есть стандартная неопределенность. Относительная стандартная неопределенность равна  $0,0577/10 = 0,006$ .

3. Часто бывает, что интервал, в пределах которого лежит измеряемая величина, известен, и мы знаем также (качественно), что измеряемая величина лежит ближе к центру, чем к краям интервала. В таком случае предполагают *треугольное распределение* случайной величины в этом интервале (см. разд. 2.3.1.3). Так обстоит дело, в частности, с дозирующими устройствами — пипетками, мерными колбами и цилиндрами.

**Пример.** Необходимо оценить неопределенность, определяемую неточностью объема мерной колбы при приготовлении водного раствора. Используется мерная колба класса А объемом 50 мл с погрешностью 0,06 мл. Известно, что при калибровке мерных колб целевое значение 50,00 мл скорее лежит ближе к середине интервала, определяемого погрешностью, чем к его краям. Поэтому принимаем случайную величину имеющей треугольное распределение в интервале  $(50,00 \pm 0,06)$  мл. Ее дисперсия равна  $s^2 = \Delta^2/6 = (0,06)^2/6 = 0,0006 \text{ (мл)}^2$  (см. разд. 2.3.1.3), стандартное отклонение (оно же стандартная неопределенность)  $s = \sqrt{0,0006} = 0,0245 \approx 0,024$  мл, а относительное стандартное отклонение  $RSD = s / \bar{X} = 0,0245/50 \approx 0,0005$  (относительная стандартная неопределенность).

4. Если достоверно выявлена постоянная систематическая погрешность  $\Delta_c$ , возможны следующие варианты действий:

- при расчете результатов измерений вводить соответствующую поправку. Если такая поправка вносится, при оценке неопределенности результата измерений необходимо учесть неопределенность поправки.
- не вносить поправку на систематическую погрешность, просто указывая при представлении результатов оценку величины  $\Delta_c$  и ее неопределенность — пусть заказчик сам разбирается;
- заменить систематическую погрешность на случайную. Для этого вместо систематической погрешности  $\Delta_c$ , не внося поправки в результат измерений, при расчете суммарной неопределенности добавляют в качестве стандартную неопределенность  $u_c(X) = k\Delta_c$ . Другими словами, в суммарную неопределенность вносится псевдослучайная составляющая, заведомо «поглощающая» систематическую погрешность. Величина коэффициента  $k$  выбирается в диапазоне между 1 и 2 и зависит от неопределенности оценки величины  $\Delta_c$ . Чаще всего  $k = 1,4$ – $1,7$ .

Последний вариант очень удобен. Конечно, при его применении могут получиться слегка завышенные оценки суммарной неопределенности, зато впоследствии не будет дополнительных операций, связанных с учетом систематической погрешности при каждом измерении.

Если после расчета стандартных отклонений какие-то из них окажутся малыми в соответствии с принципом пренебрежения малыми погрешностями (см. разд. 2.6.1), их можно не учитывать в дальнейшем.

Рассмотрим сравнительно простой пример — оценку неопределенности результатов определения концентрации изобутилметилфосфоната (*i*-БМФ), одного из алкилметилфосфонатов (АМФ), как основного вещества в стандартных образцах (то есть измеряются содержания *i*-БМФ, близкие к 100%), методом автоматического потенциометрического



титрования [85]. Метод состоит в титровании раствора исследуемого вещества раствором щелочи с помощью автоматического титратора.

Методику можно разделить на следующие этапы:

- приготовление титранта (раствор NaOH в деионизированной воде) с контролем концентрации получившегося раствора;
- взятие навески *i*-БМФ;
- собственно титрование *i*-БМФ раствором NaOH;
- расчет результата.

Расчет результатов измерений проводят по формуле  $M_d(i\text{-БМФ}) = c_{\text{NaOH}} V_T M_{\text{АМФ}} / 1000 m_{\text{АМФ}}$ , где  $M_d(i\text{-БМФ})$  – массовая доля *i*-БМФ,  $c_{\text{NaOH}}$  – концентрация раствора NaOH,  $V_T$  – объем этого раствора, пошедший на титрование АМФ,  $M_{\text{АМФ}}$  – молярная масса АМФ (в данном случае *i*-БМФ),  $m_{\text{АМФ}}$  – масса анализируемого АМФ (*i*-БМФ).

Для удобства расчета возможные компоненты неопределенности представляют в форме диаграммы «причина-следствие» (диаграмма Исикавы или стрелочная диаграмма, иногда называемой также диаграммой типа «рыбья кость», «рыбий скелет» или «еловая шишка» из-за характерной формы). Для нашего примера она показана на рис. 3.10. Построим бюджет неопределенностей.

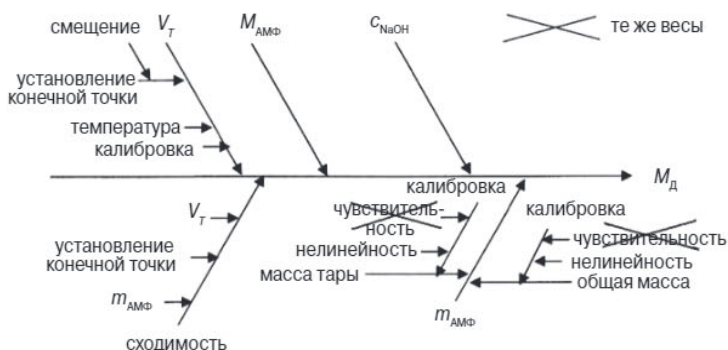


Рис. 3.10. Диаграмма «причина-следствие» (диаграмма Исикавы, стрелочная диаграмма) для рассмотренного примера – определения концентрации изобутилметилфосфоната (*i*-БМФ) методом потенциометрического титрования.  $V_T$  – объем,  $M_{\text{АМФ}}$  – молярная масса *i*-БМФ, относящегося к алкилметилфосфонатам,  $c_{\text{NaOH}}$  – концентрация NaOH,  $m_{\text{АМФ}}$  – масса *i*-БМФ,  $M_d$  – искомая концентрация (массовая доля) *i*-БМФ

1. **Сходимость.** Стандартное отклонение повторяемости (сходимости) проще всего оценить экспериментально в ходе многократного титрования одного и того же раствора в одних условиях. В данном случае  $RSD_r = 0,30\%$  ( $N = 10$ ) Эта величина включает в себя случайные погрешности всех этапов.
2.  **$m_{\text{АМФ}}$**  (в нашем примере  $m_{i\text{-БМФ}}$ ). Неопределенность взвешивания можно оценить следующим образом. Поскольку масса навески находится как разность двух близких масс (колбы с веществом и без него), систематической погрешностью весов можно пренебречь. Остается учесть погрешность градуировки весов, равную (согласно паспорту весов) для использованных весов «AND» NM-200 0,15 мг. В данном случае по рекомендации производителя весов надо использовать равномерное распределение и вклад нелинейности в каждую массу равен  $0,15 / \sqrt{3} = 0,09$  мг. Считаем оба взвешивания независимыми измерениями. В этом случае дисперсия разности равна сумме дисперсий (см. разд. 2.6.1), и стандартное отклонение (оно же неопределенность) взвешивания  $u(m) = \sqrt{2(0,09)^2} = 0,13$  мг.

3.  $V_T$ . Неопределенность объема раствора, идущего на титрование определяется (кроме уже рассмотренной сходимости):
- погрешностью автоматического титратора АТП-02, стандартное отклонение для которой была оценена как 0,005 мл;
  - неопределенностью  $u(V(t_0))$ , обусловленной возможными колебаниями температуры. Ее можно оценить по формуле  $u(V(t_0)) = (V \cdot 2,1 \cdot 10^{-4} \cdot \Delta t_0) / 1,96$ , где  $V$  – объем,  $2,1 \cdot 10^{-4}$  – коэффициент объемного расширения воды,  $\Delta t_0$  – диапазон возможных колебаний температуры раствора, 1,96 – 95% квантиль распределения Стьюдента при  $N = \infty$ ). Полагая  $\Delta t_0 = \pm 3^\circ \text{C}$ , получаем  $u(V(t_0)) = V \cdot 3,2 \cdot 10^{-4}$ .
- Смещение при установлении конечной точки можно не учитывать. Объем, идущий на титрование, практически всегда близок к 5 мл. Получаем  $u(V(t_0)) \cong 0,002$  мл и  $u(V) = \sqrt{0,005^2 + 0,002^2} \cong 0,006$  мл.
4.  $M_{\text{амф}}$  (в нашем примере  $M_{i\text{-БМФ}}$ ). Легко показать, что в соответствии с принципом пренебрежения малыми погрешностями (см. разд. 2.6.1) неопределенностью оценки молекулярной массы как  $i$ -БМФ, так и других соединений в данном случае можно пренебречь.
5. Оценим неопределенность результата одного из измерений. Необходимые данные приведены в табл. 3.4.

Таблица 3.4. Данные для расчета неопределенности

Обозначение	Наименование	Значение $X$	Стандартная неопределенность $u(X)$	Относительная стандартная неопределенность $u(X)/X$
сход	Сходимость	1,000	0,0030	0,0030
$c_{\text{NaOH}}$	Молярная концентрация NaOH в растворе	0,099492 н	0,000207 н	0,002084
$V_T$	Объем раствора NaOH, пошедший на титрование $i$ -БМФ	5,007 мл	0,006 мл	0,001662
$m_{i\text{-БМФ}}$	Масса $i$ -БМФ	0,0782 г	0,00013 г	0,001662

Результат измерения равен  $M(i\text{-БМФ}) = (0,99492 \cdot 5,007 \cdot 152,13) / (1000 \cdot 0,0782) = 96,91\%$  масс. Суммарная стандартная неопределенность (относительная) рассчитывается в данном случае по формуле (2.55):

$$\frac{u(M_a(i\text{-БМФ}))}{M_a(i\text{-БМФ})} = \sqrt{\left(\frac{u(\text{сход})}{\text{сход}}\right)^2 + \left(\frac{u(c_{\text{NaOH}})}{c_{\text{NaOH}}}\right)^2 + \left(\frac{u(V_T)}{V_T}\right)^2 + \left(\frac{u(m_{i\text{-БМФ}})}{m_{i\text{-БМФ}}}\right)^2} =$$

$$= \sqrt{0,0030^2 + 0,001662^2 + 0,002084^2 + 0,001198^2} = 0,004188.$$

Расширенная неопределенность при коэффициенте охвата 2 составляет

$$\frac{u(M_a(i\text{-БМФ}))}{M_a(i\text{-БМФ})} \cdot M_a(i\text{-БМФ}) \cdot 2 = 0,004188 \cdot 96,91 \cdot 2 = 0,8118\% \text{ масс.}$$

Округляя (см. разд. 3.5.2.4), результат измерения можно записать как  $(96,9 \pm 0,8)\%$  масс.

Рассмотренный способ оценки неопределенности, казалось бы, принципиально не отличается от оценки по известным метрологическим характеристикам методики



(разд. 3.5.2.3.1) — основные составляющие неопределенности должны в сумме давать тот же вклад, что и экспериментально оцененная воспроизводимость. Однако это не всегда так — по нашему опыту, она часто бывает завышенной, так что экспериментальное нахождение воспроизводимости (см. разд. 3.4.3) остается актуальным.

#### 3.5.2.4. Представление результатов измерений при передаче их заказчику

Как сказано выше, результат измерения, как правило, должен сопровождаться указанием его суммарной погрешности (неопределенности). Это можно делать двумя основными способами:

- указать измеренное значение и, после слов «при стандартной неопределенности», стандартную неопределенность;
- в форме (измеренное значение  $\pm$  расширенная неопределенность). При этом рекомендуется указать коэффициент охвата.

Второй способ более нагляден (см. разд. 3.5.2.3.1). Мы его используем сами и рекомендуем всем как оптимальный.

##### 3.5.2.4.1. Представление результатов измерений и их неопределенностей

Для того чтобы правильно выпisać результат измерения, надо ответить на следующие вопросы:

- сколько значащих цифр необходимо оставлять в окончательно приводимом результате измерения и в его неопределенности;
- как именно проводить округление.

С помощью современных устройств (компьютеры, калькуляторы) результат измерения и его неопределенность рассчитываются с большим числом цифр, что создает иллюзию большой точности. Однако сохранение избыточного числа значащих цифр в окончательно приводимом результате вредно, поскольку вводит в заблуждение как исследователя, так и заказчика измерений. Как же надо выписывать результаты для Заказчика?

Вначале необходимо рассчитать результат измерения (окончательно приводимое значение, см. разд. 4.4) и расширенную неопределенность с избыточным числом значащих цифр. Далее надо посмотреть, какова первая значащая цифра в расширенной неопределенности. Если это 1 или 2, то округление расширенной неопределенности проводят до двух значащих цифр, если 3 и больше, то до одной. После этого результат измерений округляют таким образом, чтобы разряд последнего знака в нем совпадал с разрядом последнего знака расширенной неопределенности (после округления).

Общее правило состоит в том, что все промежуточные расчеты проводятся без округления, с сохранением максимально большого числа значащих цифр, а округление проводится только в самом конце, перед выпиской результатов. В этом случае можно использовать стандартное округление, когда число в десятичной системе округляют до  $N$ -го знака в зависимости от  $N + 1$  знака:

- если  $(N + 1)$ -й знак  $< 5$ , то  $N$ -й знак сохраняют, а  $N + 1$  и все последующие обнуляют;
- если  $(N + 1)$ -й знак  $\geq 5$ , то  $N$ -й знак увеличивают на единицу, а  $N + 1$  и все последующие обнуляют.

Особая ситуация возникает, когда  $(N + 1)$ -й знак равен 5, а последующие знаки равны нулю. В этом случае, если округление однократное, можно использовать *математическое округление* — всегда в большую сторону. Если же по каким-то причинам округление в процессе расчетов выполняется несколько раз — лучше использовать *случайное округление* — в меньшую или большую сторону в случайном порядке, но с равной вероятностью.

Конечно, когда  $(N + 1)$ -й знак не равен 5 или последующие знаки не равны нулю, применяют стандартное округление.

Примеры того, как надо выписывать результаты измерений, даны в табл. 3.5 (использовано математическое округление).

Таблица 3.5. Примеры корректного выписывания результатов измерений

Результаты расчета в формате (результат измерения $\pm$ расширенная неопределенность)	
Без округления	Окончательно приводимый результат после округлений
$(7,1622 \pm 0,0214)$	$(7,162 \pm 0,021)$
$(35,64 \pm 0,36)$	$(35,6 \pm 0,4)$
$(846,332 \pm 8,500)$	$(846 \pm 9)$
$(0,038455 \pm 0,01493)$	$(0,038 \pm 0,015)$

#### 3.5.2.4.2. Оценка качества объекта исследований при наличии норматива

Оценка качества при наличии нормативов состоит в ответе на вопросы:

- соответствует ли содержание аналита в образце диапазону допустимых значений для данного сорта продукции (на производстве);
- не превышена ли предельно допустимая концентрация (ПДК) в объекте окружающей среды — в воде, почве и т. д. (при экологических исследованиях);
- лежит ли концентрация жизненно важных веществ — холестерина, глюкозы и т. д. — в сыворотке крови больного в диапазоне нормальных значений (в медицине) и т. д.

Такая оценка, строго говоря, не относится к метрологии химического анализа. Однако мы кратко рассмотрим этот вопрос как в силу его важности вообще, так и потому, что де-факто сотрудники лабораторий нередко в той или иной форме участвуют в подобных оценках.

На рис. 3.11 показан случай двухстороннего норматива, когда заданы как верхняя, так и нижняя границы нормы,  $A_1$  и  $A_2$ , соответственно. Бывают также односторонние нормативы: либо только верхний  $A_1$ , либо только нижний  $A_2$ <sup>63,64</sup>.

<sup>63</sup> Нормативы могут задаваться и называться по-разному: требования спецификации (на продукцию), референтные интервалы или нормальные значения (в медицине), предельно допустимые концентрации (в экологии) и т. д. В любом случае это одно — или двухсторонний интервал, в котором должна находиться концентрация аналита. Мы будем использовать выражения «норматив», «диапазон нормальных концентраций (или значений)».

<sup>64</sup> Мы будем предполагать, что в случае односторонней границы неявно существует и вторая. Так, в примерах 2 и 3 содержание кислорода лежит в диапазоне 99,999–100,000%, а содержание свинца в мясе — 0–0,5 мг/кг.

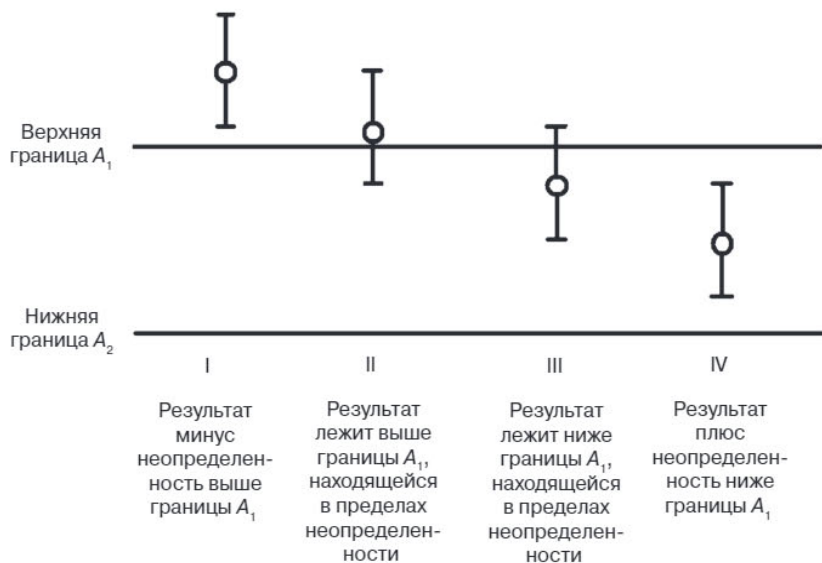


Рис. 3.11. Проверка соответствия объекта исследования нормативам. Задан двухсторонний норматив, но показаны только случаи, возможные при проверке превышения верхней границы  $A_1$

#### Примеры.

1. Концентрация глюкозы в крови взрослого человека по критериям Всемирной организации здравоохранения (ВОЗ) должна лежать в диапазоне 3,9–6,4 ммоль/л (двухсторонний норматив, заданы и нижняя граница  $A_2$  и верхняя  $A_1$ ). Если измеренная концентрация глюкозы лежит вне этого диапазона, это должно вызывать тревогу.

2. Содержание  $O_2$  в газообразном кислороде особой чистоты по ТУ 2114-001-05798345-2007 должно быть не менее 99,999% (односторонний норматив; задана только нижняя граница диапазона  $A_2$ ).

3. Согласно санитарно-эпидемиологическим правилам и нормам СанПиН 2.3.2.1078-01 содержание свинца в мясе не должно превышать 0,5 мг/кг (односторонний норматив; задана только верхняя граница диапазона  $A_1$ ).

При принятии решения о том, соответствует ли результат измерения нормативам, прежде всего необходимо разобраться, была ли учтена неопределенность измерения при установлении норматива. Если да, то надо просто сравнить окончательно приводимый результат измерения с соответствующей границей. Однако чаще всего при установлении границ диапазона нормальных концентраций неопределенность измерения не закладывается — в частности, потому, что для разных методов измерений и в разных условиях неопределенность измерения одной и той же величины может сильно различаться. *Ниже предполагается, что при установлении диапазона нормальных значений неопределенность измерений не учитывалась — границы установлены для собственно измеряемой величины (математического ожидания результатов измерений).*

Важно, что для принятия решения относительно соответствия продукции нормативам необходимы достаточно точные методы измерений. Обычно считают, что предел воспроизводимости метода  $R$  должен соответствовать требованию



$$R < 1/4(A_1 - A_2)$$

для двухстороннего норматива и

$$R < 1/2A_1 \text{ или } R < 1/2A_2$$

для односторонних нормативов. Если эти соотношения не выполняются, надо искать более точный метод.

Различные случаи, возможные при проверке соответствия объекта исследования нормативам, показаны на рис. 3.11, на котором, при заданном двухстороннем нормативе, показаны только случаи, возможные при проверке превышения верхней границы  $A_1$ . Из рисунка ясно, что для ответа на вопрос о соответствии объекта нормативу надо проверить гипотезу о равенстве результата измерений этому нормативу как точно известному значению. Такую проверку в данном случае удобно проводить с использованием интервала, в котором (с учетом неопределенности) лежит истинное значение: если интервал «накрывает» норматив, то гипотеза не отвергается (см. разд. 2.1.4, второй способ проверки гипотез). Обозначим этот интервал как  $(X \pm U)$ , где  $U = ku_c(X)$  – расширенная неопределенность результата измерения  $X$  с коэффициентом охвата  $k$ .

Если гипотеза отвергается с принятой доверительной вероятностью, то есть граница  $A_1$  лежит вне интервала  $(X \pm U)$ , то можно сделать однозначный вывод, а именно:

- если результат измерения  $X$  лежит выше верхней границы  $(X \pm U)$  – случай (i), то норматив достоверно превышен (в случае промышленной продукции это брак, в случае объектов окружающей среды достоверное превышение ПДК и т. д.);
- если результат измерения  $X$  лежит ниже верхней границы  $A_1$  – случай (iv), то измеряемая величина не превышает норматив (качественная продукция, допустимое содержание вредного вещества ниже ПДК и т. д.).

Если гипотеза не отвергается, то есть интервал  $(X \pm U)$  «накрывает» границу  $A_1$  (случаи (ii) и (iii) на рис. 3.11), то однозначных выводов сделать нельзя. Считается, что в таком случае необходимо применить более точную методику измерений.

Аналогично рассматривается и вопрос о том, лежит ли измеряемая величина ниже нижней границы  $A_2$ .

В простейшем случае неопределенность  $U$  можно рассчитать, полагая распределение результатов нормальным, а воспроизводимость метода  $s_R$  известной. Из свойств нормального распределения (см. разд. 2.3.1.1) известно, что в этом случае при доверительной вероятности  $\bar{P} = 0,95$  доверительный интервал для значения  $X$  есть  $(X \pm 1,64s_R)$ . Поскольку  $R = 2,77s_R$ , получаем  $U = 0,59R$  и интервал  $(X \pm 0,59R)$  при коэффициенте охвата 1,96.

Существует и более простой подход, при котором измеренная величина просто сравнивается с нормативом, без учета доверительных интервалов. Другими словами, если измеренная величина  $X$  лежит выше верхней границы  $A_1$  или ниже нижней границы  $A_2$ , то это брак, а если наоборот – качественное изделие. В этом случае вероятность ошибки при интерпретации результатов очень велика, но риск ошибки делится пополам между поставщиком и потребителем.



### 3.6. Межлабораторные эксперименты

*Под межлабораторным экспериментом<sup>65</sup> понимают эксперимент, организованный с целью выявления и объективной оценки межлабораторных (а зачастую и внутрилабораторных) погрешностей измерения [86—88]. Межлабораторным экспериментом являются, в частности, кратко рассмотренные выше эксперименты по аттестации стандартных образцов с участием нескольких лабораторий (см. разд. 3.2.4), исследования метрологических характеристик методик КХА (разд. 3.4.3), исследования, проводимые для оценки состояния измерения в той или иной сфере, а также программы внешней оценки качества (см. разд. 4.6).*

Межлабораторный эксперимент организуется следующим образом. В несколько лабораторий одного профиля рассылают один или несколько *контрольных материалов* (как и прежде, мы будем использовать этот термин для обозначения однородных, стабильных и адекватных материалов, которые до эксперимента могут либо иметь, либо не иметь опорные значения), состав которых этим лабораториям неизвестен. При этом рассылаемые материалы могут быть аттестованными (стандартные образцы, образцы сравнения) или нет, но обязательно однородными и стабильными на протяжении эксперимента. В ходе межлабораторных экспериментов практически всегда удастся оценить концентрацию аналитов в исследуемых материалах. Схема анализа материалов (дизайн межлабораторного эксперимента) задается его организаторами.

Вообще говоря, в начале реализации межлабораторного эксперимента должен составляться *протокол* — документ, содержащий полную информацию об этом эксперименте. Особое значение протокол имеет при внешней оценке качества, поэтому требования к этому документу перечислены нами в разд. 4.6.

Для достижения большей объективности организаторы межлабораторного эксперимента часто прибегают к разным ухищрениям: кодируют контрольные материалы, рассылают образцы с известными только им добавками, учитываемыми при обработке данных, ограничивают массу материала и т. д. Это особенно важно для правильной оценки внутрилабораторных погрешностей, без достоверного знания которых ценность межлабораторного эксперимента резко снижается.

#### 3.6.1. Контрольные материалы и дизайн межлабораторного эксперимента

К однотипным<sup>66</sup> материалам, рассылаемым в лаборатории в ходе межлабораторного эксперимента, предъявляются обычные для образцов сравнения требования, рассмотренные выше (разд. 3.2).

<sup>65</sup> В англоязычной литературе используется несколько терминов, обозначающих межлабораторный эксперимент вообще или конкретные его разновидности: *interlaboratory test*, *interlaboratory comparison*, *collaborative study*, *collaborative trial*, *co-operative trial* и т. д.

<sup>66</sup> Здесь и ниже *однотипными* мы называем материалы с близким составом матрицы и похожими физико-химическими свойствами, которые могут отличаться по концентрации аналитов. Строго говоря, различия в составе и свойствах должны быть таковы, чтобы ими можно было пренебречь с точки зрения влияния на результаты анализа. Можно сказать, что однотипные материалы должны быть адекватны друг другу.

Для интерпретации результатов межлабораторного эксперимента нужна информация о концентрации определяемых веществ (аналитов) в контрольных материалах. Способы оценки этой концентрации не отличаются от описанных в разд. 3.3; чаще всего это делается в ходе самого межлабораторного эксперимента.

В каждой лаборатории присланные контрольные материалы анализируются по заданной организатором эксперимента — *координатором* — схеме (план эксперимента, дизайн). Вообще говоря, материалы могут исследоваться в нескольких аналитических сериях, с выполнением нескольких измерений в каждой серии. Такая структура позволяет оценивать как собственно межлабораторные, так и внутрилабораторные погрешности (последние могут быть поделены на меж- и внутрисерийные). Программы внешней оценки качества, проводимые по подобным схемам, часто называют «обучающими».

На практике по экономическим и другим причинам чаще используют упрощенные планы межлабораторных экспериментов. Так, исключается требование выполнения анализов в нескольких сериях. В простейшем варианте в эксперименте исследуют один контрольный материал, и одна лаборатория представляет один результат (это схема, обычно применяемая при аттестации стандартных образцов и часто реализуемая в программах внешней оценки качества). Информация, которую можно получить при использовании упрощенных схем, ограничена. Особенно негативно сказывается отсутствие нескольких, хотя бы двух измерений одного материала в каждой лаборатории: в этом случае затрудняется применение многих статистических методов анализа результатов (см. ниже).

Координатор межлабораторного эксперимента обычно интересуется используемыми методами (методиками) анализа, аппаратурой и т. д. Эти сведения должны учитываться при статистическом анализе и интерпретации результатов.

### *3.6.2. Способы, применяемые для получения объективной информации о качестве анализа*

На первый взгляд кажется, что в получении объективной информации нет никаких проблем: достаточно разослать контрольные материалы и задать план их исследования. Однако все не так просто.

Во многих случаях руководители и сотрудники лабораторий могут быть заинтересованы в представлении «хороших» результатов, чтобы выглядеть наилучшим образом. Это может повысить рейтинг лаборатории и, как следствие, улучшить их материальное положение. Конкретные сотрудники, выполняющие анализ, также заинтересованы в «хороших результатах», свидетельствующих об их высокой квалификации.

Для получения объективных оценок координаторы вынуждены принимать специальные меры. В некотором смысле их задача схожа с организацией экзаменов в учебных заведениях — необходимо исключить любую возможность фальсификации результатов.

В зависимости от специфики деятельности лабораторий, используемых образцов сравнения и дизайна эксперимента, для достижения этой цели используют различные приемы:



- соблюдение анонимности лабораторий (конфиденциальность);
- исследование контрольных материалов под видом обычной (рутинной) пробы;
- кодирование контрольных материалов;
- внесение дополнительного количества аналитов в контрольные материалы;
- требование одновременности выполнения анализов;
- ограничение массы контрольных материалов, передаваемых в лабораторию;
- контроль первичных результатов измерений (приборных распечаток и пр.);
- умышленные несущественные нарушения дизайна и т. д.

Рассмотрим эти «маленькие хитрости» подробнее.

**Анонимность лабораторий** — важнейший принцип организации межлабораторных экспериментов. Каждая лаборатория участвует в эксперименте под условным номером, известным лишь лаборатории и координатору. Полученные результаты в публикуемых работах могут соотноситься только с этим номером, но не с названием лаборатории. Координаторы должны принимать меры по неразглашению конфиденциальной информации.

Гарантированная анонимность упрощает вовлечение лабораторий в межлабораторные эксперименты. В целом руководители лабораторий обычно заинтересованы в участии для получения объективной информации о качестве анализа в лабораториях. Однако опасения оказаться в числе «плохих» могут сдерживать их или заставить стремиться к «улучшению» результатов, анонимность же лишает эти опасения оснований и в конечном счете делает результаты эксперимента более объективными. С другой стороны, каждая лаборатория является собственником полученных ею результатов и может их публиковать.

**Исследование контрольных материалов под видом обычной (рутинной) пробы** — идеальный способ получения объективной информации. Тот фактор, что никто в лаборатории (или, по крайней мере, исполнители анализа) не знает, что данный образец — контрольный материал, безусловно, содействует получению объективной информации о качестве анализа в лаборатории.

К сожалению, такой прием редко удается реализовать. Это связано с тем, что в большинстве случаев образцы сравнения сильно отличаются от рутинных проб по внешнему виду, упаковке и т. д. Кроме того, в принципе создать ситуацию, в которой исполнители не знают, что исследуют контрольный материал, удается очень редко. Относительно просто сделать это в случае, когда лаборатория постоянно анализирует пробы, поступающие в лабораторию уже закодированными (например, при некоторых видах арбитражного анализа).

**Кодирование контрольных материалов** — наиболее распространенный прием, используемый в большинстве межлабораторных экспериментов. Каждый экземпляр материала, направляемый в лабораторию, кодируется, причем способ кодирования известен только организаторам эксперимента. Если лаборатория должна выполнить несколько измерений для каждого материала, целесообразно выслать несколько экземпляров каждого из них и закодировать каждый экземпляр отдельно.

Эффективность кодирования снижается в тех случаях, когда в лаборатории рассылается небольшое число контрольных материалов, сильно различающихся внешним видом или концентрацией определяемых веществ.

*Внесение дополнительного количества аналитов в контрольные материалы* заключается в том, что материалы готовятся в два этапа. На первом этапе материал готовят обычным образом. На втором этапе в каждый экземпляр вносят некоторое точно известное количество аналита (аналитов), соизмеримое с их исходной концентрацией. При обработке результатов измерений из них предварительно вычитают значения добавленных концентраций.

Такой прием очень эффективен, однако его использование возможно лишь при выполнении условий применения метода введено-найденно или метода добавок (см. разд. 3.4.4).

Требование *одновременности выполнения анализов* легко выполнимо. Обычно лаборатории просят выполнить измерения в течение одного и того же короткого промежутка времени (например, в течение недели) — это всегда полезно в случаях, когда нет проблем с доставкой материалов в лаборатории к заранее определенному сроку. Особые преимущества такой прием дает применительно к не вполне стабильным материалам — изменения в их составе происходят примерно одинаково в разных лабораториях (при этом желательно унифицировать условия хранения контрольных материалов в лабораториях и при транспортировке).

*Ограничение массы контрольных материалов, передаваемых в лабораторию*, связано с тем, что лаборатории часто стараются выполнить больше измерений, чем требуется по плану эксперимента, и представляют координаторам усредненные результаты. При этом занижаются как повторяемость (сходимость), так и общая погрешность. Для исключения этого координаторы по результатам анкетирования или путем экспертной оценки определяют количество вещества, минимально необходимое для выполнения анализов в соответствии с выбранным планом эксперимента. Такое количество контрольных материалов и высылается в лаборатории. Оно может быть разным для разных лабораторий.

*Контроль первичных результатов измерений (например, приборных распечаток)* обычно возможен при использовании автоматизированных систем измерения. Проверка их координатором затрудняет «улучшение» результатов анализа. При этом важна сама возможность контроля — требование хранить распечатки для возможной инспекции, а реальные проверки могут проводиться выборочно.

*Умышленное нарушение дизайна* может осуществляться разными способами. Например, согласно плану эксперимента в лаборатории должен рассылаться один и тот же контрольный материал, однако в одну из лабораторий высылается другой, который был аттестован ранее. Основное требование к таким нарушениям дизайна — они не должны существенно снижать информативность эксперимента.

Возможны и другие приемы, повышающие объективность оценок в рамках межлабораторных экспериментов. Многое зависит от фантазии и финансовых возможностей координаторов.

### 3.6.3. Методы обработки результатов межлабораторных экспериментов

Ниже рассмотрен процесс обработки результатов, проводимый индивидуально по каждому определяемому веществу.



### 3.6.3.1. Схемы с одним контрольным материалом

**Простейший дизайн.** В этом случае в рамках межлабораторного эксперимента исследуется один контрольный материал, причем каждая лаборатория выполняет только одно измерение и, соответственно, представляет один результат. Это обычная схема, используемая, например, при установлении содержания аналитов в стандартных образцах путем межлабораторного эксперимента (см. выше, разд. 3.2.4). Иногда при этом лаборатория представляет и свою оценку доверительного интервала для полученного ею результата анализа. Такая структура эксперимента предоставляет весьма ограниченные возможности для статистической обработки результатов. Фактически можно изучать лишь распределение результатов измерений.

Для распределений, близких к нормальному, оценивают характеристики распределения результатов (среднее значение, стандартное отклонение и т. д.) после исключения выпадающих значений (см. разд. 2.4.2). Полученное при этом среднее значение есть аттестованное в рамках эксперимента значение для данного образца. Если в эксперименте использовался ранее аттестованный образец, то в рамках межлабораторного эксперимента проверяют гипотезу о равенстве двух математических ожиданий (см. разд. 2.5.3). Часто находят долю от общего числа участвующих в эксперименте лабораторий, удовлетворяющую тем или иным критериям (например, лежащим в определенном интервале вблизи среднего значения). Таким образом удается сделать самые общие выводы о качестве работы лабораторий.

Дополнительную информацию получают на основании сведений о лабораториях (используемые методы анализа и аппаратура, наличие или отсутствие аккредитации и т. д.). Однако при этом возможны ограничения, обусловленные упрощенным дизайном эксперимента, который не обеспечивает достоверную информацию о внутрилабораторной погрешности измерений (с точки зрения статистики — об остаточной дисперсии). И даже если лаборатории сообщают доверительный интервал для полученных ими результатов, эти сведения, как правило, недостоверны из-за субъективности и различия способов оценки этих интервалов. Для корректной оценки внутрилабораторной дисперсии и обеспечения возможности применения дисперсионного анализа используют схемы межлабораторного эксперимента с повторными измерениями в лабораториях.

**Схема с одним образцом и повторными измерениями** предусматривает, что среди лабораторий распространяют один контрольный материал, который анализируется в каждой лаборатории два или более раз. В результатах такого эксперимента содержится вся информация, что и в случае простейшего дизайна (для ее получения достаточно работать со средними значениями, полученными в каждой лаборатории). Однако проведение параллельных измерений дает новые возможности.

При работе по этой схеме возможны два основных варианта: 1) повторные измерения равноправны, т. е. проводятся либо в одной серии анализа (рис. 3.12а), либо по одному в каждой из нескольких серий (рис. 3.12б) и 2) измерения проводятся в нескольких сериях анализа, по два или более измерения в аналитической серии (рис. 3.12в).

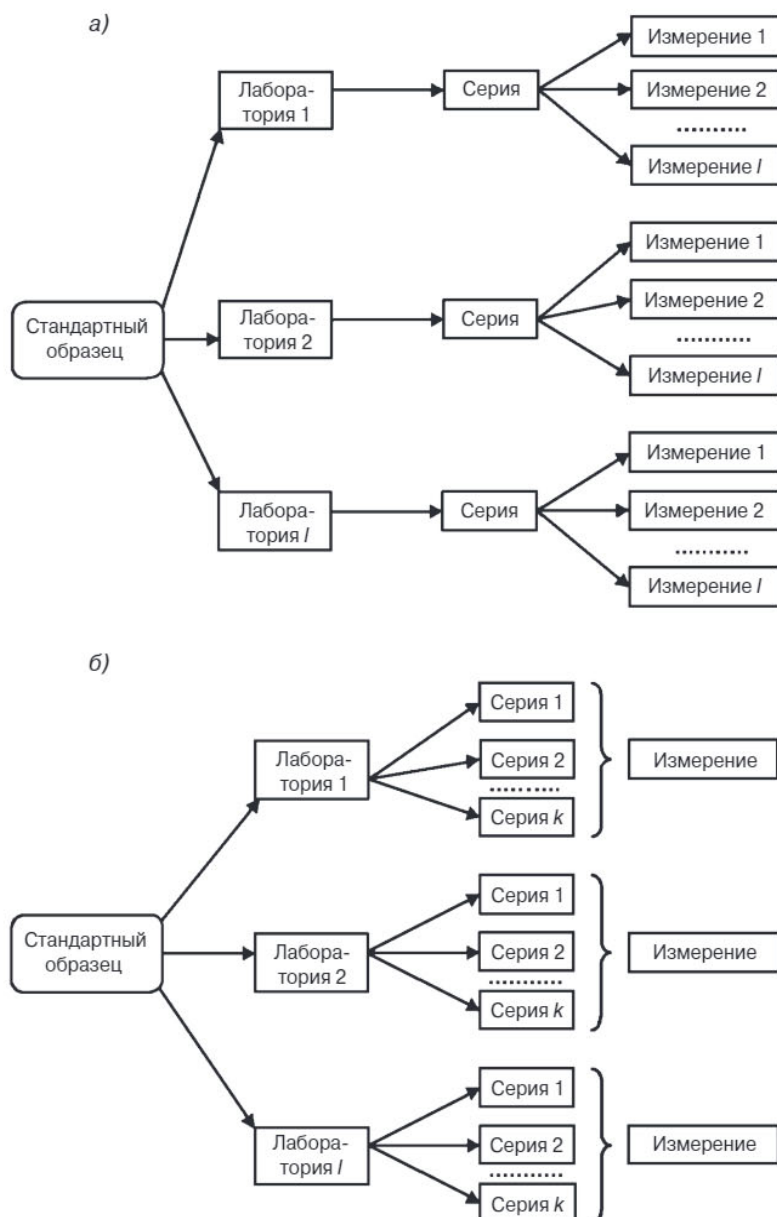


Рис. 3.12. Планы межлабораторного эксперимента с одним исследуемым образцом: равноправные измерения (а, б)

В случае равноправных измерений, т. е. либо проведенных в одной серии, либо по одному измерению в серии (рис. 3.12а, б), результаты могут быть рассмотрены в рамках модели

$$Y_{ij} = \alpha + \alpha_i + \varepsilon_{ij}, \quad (3.42)$$

в)

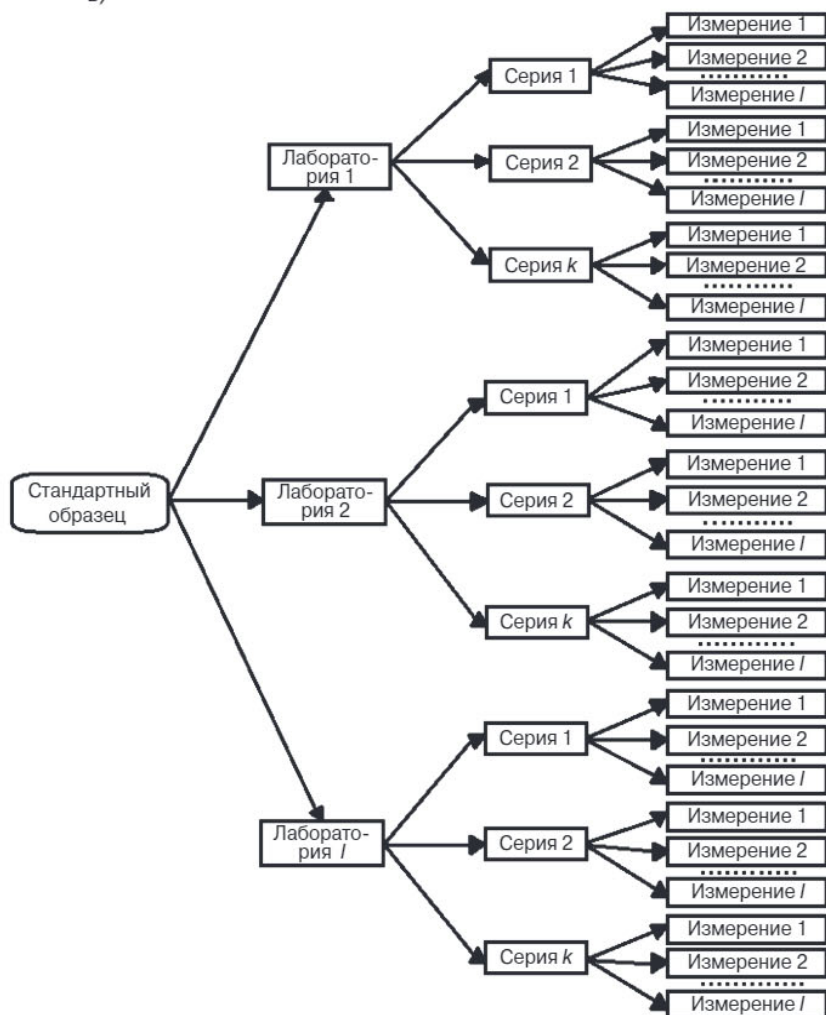


Рис. 3.12 (окончание). Планы межлабораторного эксперимента с одним исследуемым образцом: «вложенная» структура (в)

где  $i = 1, 2, \dots$  — номер лаборатории,  $l = 1, 2, \dots$  — номер повторного измерения;  $Y_{il}$  — результат  $l$ -го измерения в  $i$ -й лаборатории;  $\alpha$  — среднее значение;  $\alpha_i$  — эффект лаборатории;  $\varepsilon_{il}$  — случайная погрешность.

Это — модель однофакторного дисперсионного анализа (см. разд. 2.6.2). Как обычно, предполагают нормальным распределение эффектов  $\alpha_i$  и  $\varepsilon_{il}$ .

С помощью модели (3.42) могут быть получены корректные ответы на вопросы: 1) каковы оценка среднего значения  $\alpha$  и доверительный интервал для этой оценки при заданном уровне доверительной вероятности; 2) имеется ли эффект лаборатории, т. е. отличаются ли лаборатории между собой в совокупности, для этого необходимо проверить гипотезу  $H_0: (\alpha_i = 0)$ ; 3) каковы оценки стандартного



отклонения межлабораторной и остаточной (внутрилабораторной) составляющих погрешности; их раздельная оценка имеет смысл в случае положительного ответа на предыдущий вопрос. Если расчеты проводятся для лабораторий, использующих одну аналитическую методику, то межлабораторный разброс соответствует воспроизводимости этой методики, а внутрилабораторный — повторяемости (сходимости), если измерения выполнялись в одной серии, и промежуточной прецизионности — если в разных.

При наличии дополнительной информации об используемых методах анализа и аппаратуре, об аккредитации лаборатории или отсутствии ее и др. можно исследовать влияние этих факторов в рамках модели типа (3.42), в которой  $i$  — фактор (например, метод анализа),  $k$  — номер лаборатории (см. ниже).

Если дизайн эксперимента предусматривает исследование образца в нескольких аналитических сериях с проведением по два или более измерений в серии (рис. 3.12в), то результаты могут быть рассмотрены в рамках модели:

$$Y_{ikl} = \alpha + \alpha_i + \beta_{ik} + \varepsilon_{ikl}, \quad (3.43)$$

где  $i$  — номер лаборатории;  $k$  — номер серии;  $l$  — номер повторного измерения;  $Y_{ikl}$  — результат  $l$ -го повторного измерения в  $k$ -й аналитической серии в  $i$ -й лаборатории;  $\alpha$  — среднее значение;  $\alpha_i$  — эффект лаборатории;  $\beta_{ik}$  — эффект серии;  $\varepsilon_{ikl}$  — случайная погрешность.

Это — типичная модель дисперсионного анализа с «вложенным» фактором (эффект  $\beta_{ik}$  «вложен» в эффект  $\alpha_i$ ). В рамках модели дисперсионного анализа (3.43) дополнительно можно выяснить, имеется ли в лабораториях достоверная межсерийная вариация результатов, т. е. проверить гипотезу  $H_1: (\beta_{ik} = 0)$ . Если эта гипотеза отвергается, то можно дополнительно получить оценку эффекта  $\beta_{ik}$ . С точки зрения химической метрологии в (3.43) эффект аналитической серии есть промежуточная прецизионность, а случайная погрешность — это повторяемость (сходимость).

Если гипотеза  $H_1: (\beta_{ik} = 0)$  не отвергается, статистический анализ результатов эксперимента проводится в рамках модели (3.42).

### 3.6.3.2. Схемы с несколькими контрольными материалами<sup>67</sup>

Межлабораторный эксперимент, выполняемый с использованием нескольких однотипных материалов с разным содержанием определяемых веществ (аналитов), открывает возможность изучения зависимостей погрешностей от концентрации. При этом могут быть реализованы три основные схемы организации межлабораторного эксперимента, фактически повторяющие для каждого из образцов описанные выше схемы.

**Один результат для каждого образца** — простейший дизайн, воспроизводимый на нескольких образцах сравнения. При такой схеме эксперимента можно провести статистический анализ, описанный в разд. 3.6.3.1, для каждого из образцов.

<sup>67</sup> Здесь и ниже, когда говорится о нескольких контрольных материалах, имеются в виду однотипные материалы (см. сноску 66). Если организаторы эксперимента рассылают несколько существенно разных образцов, это — отдельные эксперименты.

Ограничения, налагаемые отсутствием повторных измерений, не позволяют проверить многие гипотезы.

Другим способом обработки результатов такого эксперимента для случая исследования двух контрольных материалов являются диаграммы Юдена [22, 89].

Условия построения таких диаграмм следующие:

- обрабатываются результаты анализа двух материалов;
- пренебрегают изменением случайных погрешностей при переходе от одного материала к другому;
- результаты получаются в лабораториях одним и тем же методом.

Результаты, полученные в лабораториях, представляют графически в виде диаграмм Юдена, которые строят в координатах «концентрация определяемого вещества в одном стандартном образце—концентрация в другом стандартном образце» (рис. 3.13). По виду полученной диаграммы рассеяния судят о наличии или отсутствии систематических погрешностей, об их соотношении со случайными погрешностями и о наличии «выпадающих» лабораторий.

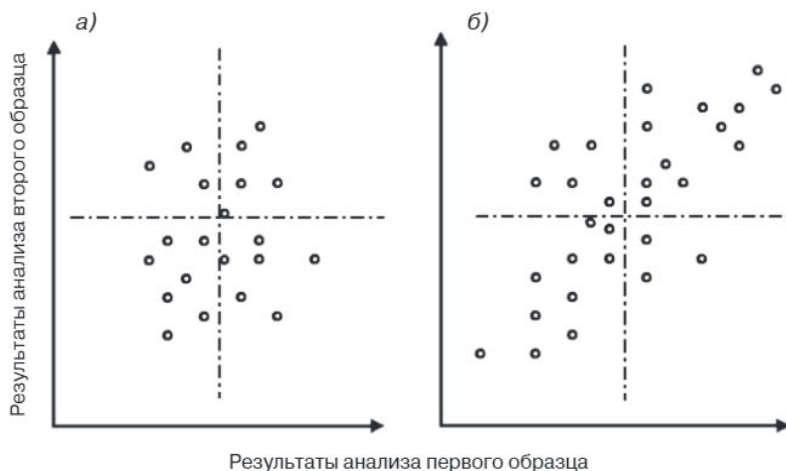


Рис. 3.13. Диаграмма Юдена для случая, когда случайные погрешности в лабораториях близки: систематическая погрешность отсутствует (а), имеется значительная систематическая погрешность (б)

Кроме того, оценки, полученные отдельно по каждому стандартному образцу, можно использовать для нахождения зависимости погрешностей от концентрации определяемых веществ. Обычно такие зависимости хорошо описываются степенной функцией

$$\sigma = \sigma_0 C^k, \quad (3.44)$$

где  $\sigma$  – стандартное отклонение, соответствующее рассматриваемому эффекту;  $C$  – концентрация;  $\sigma_0$  и  $k$  – константы.

Чаще всего зависимость погрешности от концентрации представляют в преобразованном виде

$$\lg V = -p \lg C + \lg V_0, \quad (3.45)$$

где  $V$  – относительное стандартное отклонение;  $p = 1 - k$  ( $k$  – показатель степени в формуле 3.44);  $\lg V_0$  – константа.

Зависимости вида (3.43, 3.44) могут быть найдены для погрешностей разных видов (например, для повторяемости (сходимости) и воспроизводимости) сравнения.

Примеры зависимостей вида (3.44, 3.45) для общей погрешности методик, полученные в разных областях аналитической химии, приведены в [90].

*Схема с несколькими образцами и несколькими повторными измерениями*, как и в случае экспериментов с использованием одного образца, резко расширяет возможности статистического анализа результатов межлабораторного эксперимента. Применяют две основные схемы такого типа: 1) с несколькими равноправными измерениями (т. е. проводимыми либо по одному в каждой из нескольких аналитических серий, либо в одной серии) и 2) с проведением измерений в нескольких аналитических сериях, по два или более измерения в серии. Эти схемы аналогичны показанным на рис. 3.12 с той лишь разницей, что во всех лабораториях в каждой серии осуществляется (однократно или многократно) анализ не одного, а нескольких однотипных материалов.

При реализации таких схем результаты измерений можно обрабатывать отдельно по каждому образцу с использованием моделей дисперсионного анализа (3.41) или (3.42). Кроме того, наличие данных по нескольким однотипным материалам дает возможность качественного рассмотрения зависимости получаемых оценок погрешностей от концентрации определяемых веществ. Для каждого из оцененных эффектов могут быть найдены зависимости вида (3.43, 3.44). Такие зависимости имеют самостоятельное значение, а также могут быть использованы в расчетах по моделям (3.45–3.47, см. ниже).

Результаты, полученные для нескольких однотипных контрольных материалов с повторными измерениями, можно обрабатывать в рамках моделей, учитывающих эффект концентрации.

Рассмотрим схему с равноправными измерениями. В этом случае результаты межлабораторного эксперимента можно описать моделью дисперсионного анализа со шкалированием (делением полученного результата на опорное (аттестованное) значение для соответствующего образца сравнения) [91–93]:

$$\frac{Y_{ijl}}{C_j} = \mu + \lambda_i + \varphi_{ij} + \xi_{ijl}, \quad (3.46)$$

где  $i$  – номер лаборатории;  $j$  – номер образца сравнения;  $l$  – номер повторного измерения;  $Y_{ijl}$  – результат  $l$ -го повторного измерения в  $j$ -м образце в  $i$ -й лаборатории;  $C_j$  – результат аттестации  $j$ -го образца (если образцы аттестуются в ходе эксперимента,  $C_j$  – среднее всех измерений для  $j$ -го образца после отброса выпадающих результатов);  $\mu$  – среднее значение всех величин  $Y_{ijl}/C_j$ ;  $\lambda_i$  – эффект лаборатории;  $\varphi_{ij}$  – эффект стандартного образца;  $\xi_{ijl}$  – случайная погрешность. По сути дела в рамках модели (3.46) работают с относительными смещениями  $B_{ijl} = (Y_{ijl} - C_j)/C_j$ , отличающимися от левой части модели на единицу. При расчетах обычно полагают, что эффекты  $\lambda_i$ ,  $\varphi_{ij}$  и  $\xi_{ijl}$  имеют нормальное распределение.



В наиболее частом случае, когда гипотеза  $H_0: (\mu = 1)$  не отвергается, это означает, что все погрешности рассматриваются как пропорциональные, и модель (3.45) можно записать в виде

$$B_{jil} = \frac{Y_{jil}}{C_j} - 1 = \lambda_i + \varphi_{ij} + \xi_{jil}, \quad (3.47)$$

где в левой части находятся относительные смещения  $B_{jil}$ .

Геометрическая интерпретация модели (3.46) представляет собой семейство прямых, исходящих из начала координат (рис. 3.14).

При статистическом анализе в рамках модели (3.46) вначале проверяют гипотезу  $H_0: (\varphi_{ij} = 0)$ . Если эта гипотеза отвергается, то, значит, используемые образцы сравнения скорее всего не однотипны. Проверка гипотезы  $H_1: (\lambda_i = 0)$  позволяет выявить пропорциональную погрешность, обусловленную эффектом лабораторий. Наличие достоверно отличающегося от нуля эффекта  $\lambda_i$  (особенно, если его оценка значительна по величине) чаще всего свидетельствует о погрешностях градуировки, пропорциональных погрешностях измерительной аппаратуры и т. д. Отметим, что модель (3.47) практически нечувствительна к выбору концентраций определяемых веществ в образцах.

Другая модель расчета относится к результатам, полученным с использованием нескольких образцов сравнения и равноправных измерений, — модель ковариационного анализа [92, 93]

$$Y_{jil} = \eta_i + \Psi_i C_j + \xi_{jil}, \quad (3.48)$$

где индексы и величины  $Y_{jil}$  и  $C_j$  те же, что в модели (3.45);  $\eta_i$  и  $\Psi_i$  — соответственно, отрезок, отсекаемый на оси ординат, и угловой коэффициент уравнений регрессии для результатов  $i$ -й лаборатории;  $\xi_{jil}$  — случайная ошибка.

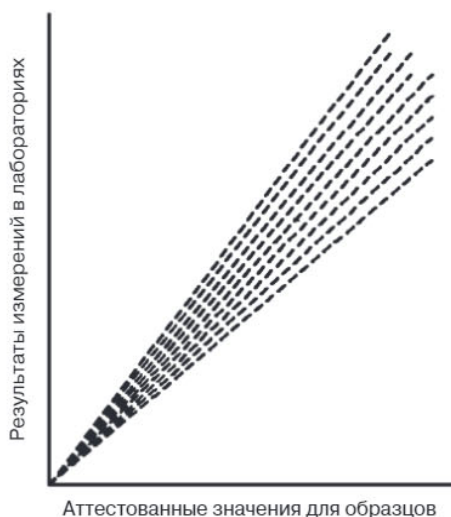


Рис. 3.14. Геометрическая интерпретация модели (3.47)



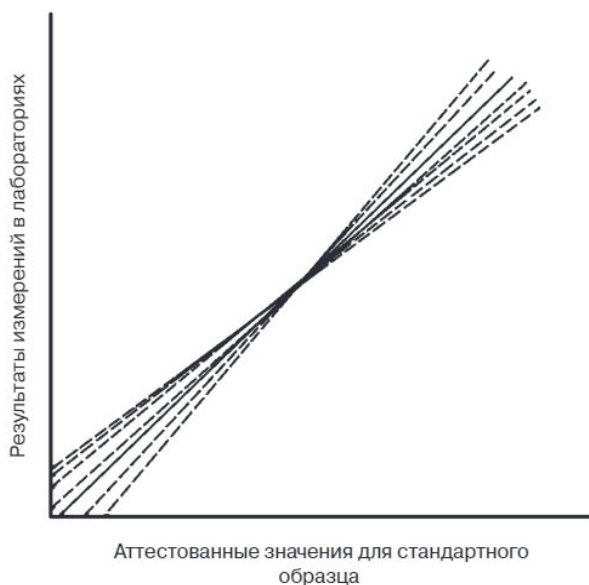


Рис. 3.15. Геометрическая интерпретация модели (3.47)

Геометрическая интерпретация для модели (3.47) — семейство прямых линий, каждая из которых отображает уравнение регрессии для результатов одной лаборатории (рис. 3.15). Это — модель ковариационного анализа [92, 93].

Во всех случаях при статистических расчетах необходимо учитывать гетероскедастичность (неоднородность дисперсий). Это можно сделать введением весовых коэффициентов  $w$  («взвешенный» метод наименьших квадратов, см. разд. 2.7.2.3). В частности, если известна зависимость вида (3.44), то  $w = 1/(V_0 C^{-p})$ .

Статистический анализ результатов в рамках модели (3.48) позволяет ответить на вопросы: 1) имеется ли достоверная вариация угловых коэффициентов уравнений регрессии между лабораториями, т. е. проверка гипотезы  $H_2(\Psi_i = 0)$ , проверка на присутствие межлабораторной погрешности, пропорциональной концентрации; 2) имеется ли достоверная вариация между отрезками, отсекаемыми на оси ординат уравнений регрессии между лабораториями, т. е. проверка гипотезы  $H_4: (\eta_i = 0)$ , проверка на присутствие постоянной, не зависящей от концентрации, погрешности между лабораториями.

Постоянная и пропорциональная составляющие погрешности обусловлены разными причинами: постоянная — главным образом некорректным учетом результатов холостого опыта, пропорциональная — это погрешности градуировки, дозирования пробы и т. д. [94]. Поэтому статистический анализ в рамках модели (3.47) позволяет не только констатировать наличие межлабораторных погрешностей того или иного вида, но и во многих случаях указать пути их снижения.

Если используются предварительно аттестованные, например, образцовым методом, образцы сравнения (т. е.  $C_j$  оценены вне данного эксперимента), можно проверить, достоверны ли общие для всех лабораторий пропорциональная и постоянная составляющие погрешности. Для этого модель (3.48) надо записать в виде

$$Y_{ijl} = \phi_0 + \phi_i + (\theta_0 + \theta_i)C_j + \xi_{ijl}, \quad (3.49)$$

где  $\phi_0 = \sum_{i=1}^I \eta_i$ ,  $\theta_0 = \sum_{i=1}^I \Psi_i$  ( $I$  – общее число лабораторий);  $\phi_i \in N(0, \sigma_\phi)$ ,  $\theta_i \in N(0, \sigma_\theta)$ , и проверить гипотезы  $H_5(\phi_0 = 0)$ ,  $H_6(\theta_0 = 1)$ . Если одна из этих гипотез (или обе) отвергаются, это свидетельствует о том, что все лаборатории в совокупности имеют постоянную или пропорциональную погрешность (или и ту и другую вместе).

Расчеты в рамках модели дисперсионного анализа (3.47) и моделей (3.48, 3.49) дополняют друг друга: возможны ситуации, когда погрешности выявляются одной из них и не выявляются другими. Заметим также, что использовать дающую много информации модель ковариационного анализа надо с осторожностью, поскольку она может быть неустойчивой при неудачно выбранном дизайне – например, если интервал концентрации контрольных материалов недостаточно широкий. В случае появления сомнений лучше проверить устойчивость модели по отношению к реальным данным, внося небольшие изменения в исходный цифровой материал. Это не должно сильно сказываться на результатах статистического анализа.

Если каждый образец анализируется по схеме «несколько раз в нескольких сериях», в моделях (3.46–3.48) появляется вложенный эффект серии аналогично тому, как это имеет место при переходе от модели (3.42) к модели (3.43). Аналогичная ситуация – появление в моделях (3.47–3.49) новых эффектов – возникает при исследовании влияния тех или иных факторов (метод анализа, тип лаборатории и т. д.) или их взаимодействий на результаты анализа. Способы оценки этих эффектов описаны в статистических руководствах (см., например, [35, 42, 47, 95]) и здесь мы не будем на них останавливаться.

Таким образом, обрабатывая результаты межлабораторного эксперимента индивидуально по каждому аналиту и изменяя дизайн эксперимента, можно получать разнообразную информацию о составляющих погрешности.

Разновидностью межлабораторных экспериментов является внешняя оценка качества (межлабораторные сравнительные испытания, Proficiency Testing). Она рассмотрена в разд. 4.6.

### 3.7. Аккредитация лабораторий

**Аккредитация** (от лат. *accredo*, «оказываю доверие»), вообще говоря, – процедура, результатом которой является официальное подтверждение соответствия качества предоставляемых услуг некоему стандарту<sup>68</sup>. Она распространена в сфере оказания профессиональных услуг, для оценки качества которых потребитель, как правило, не обладает достаточной компетентностью. Нас интересует аккредитация лабораторий, выполняющих, наряду с другими исследованиями, химический анализ (аналитических лабораторий).

<sup>68</sup> Термин «аккредитация» используется в разных областях и может иметь существенно разный смысл. В международном праве аккредитация – процедура назначения дипломатического представителя (или постоянного представителя государства при какой-либо международной организации). Аккредитации подлежат также представители средств массовой информации. Мы используем этот термин в более узком смысле.

*Аккредитация аналитической лаборатории* есть официальное признание компетентности (технической) лаборатории, осуществляющей количественный химический анализ. Под аккредитацией обычно понимают как процедуру проверки лаборатории на компетентность (процедура аккредитации), так и ее результат («аккредитованная лаборатория»).

*Компетентность (техническая)* аналитической лаборатории есть способность выполнять определенные виды химического анализа (и других измерений) с требуемой точностью. Под видами химического анализа понимается четко определенный круг объектов анализа и конкретные вещества, содержание которых в этих объектах измеряется («*область аккредитации*»). Для признания технической компетентности необходимо наличие в лаборатории оборудования (или свободного доступа к нему), реактивов и документов, необходимых для проведения количественного химического анализа, условий для измерений, а также достаточного по численности и квалификации персонала и системы обеспечения качества анализа. При аккредитации лабораторий, выполняющих рутинные исследования, обычно требуются использование ею аттестованных методик. Основные требования к лабораториям, в том числе и аналитическим, содержатся в международном стандарте [73].

При аккредитации иногда проверяют также *независимость* лаборатории, т. е. отсутствие возможности какого-либо давления извне или заинтересованности сотрудников лаборатории, что может повлиять на выдаваемые лабораторией результаты. Для признания независимости лаборатории обычно требуется наличие у нее юридического статуса, кроме того, должны быть исключены воздействия на сотрудников лаборатории, а также связь результатов анализа с вознаграждением персонала. В любом случае *лаборатория не должна участвовать в деятельности, которая может вызывать сомнения в ее объективности*.

Аккредитацию лабораторий проводят органы по аккредитации.

Аккредитация является добровольной процедурой.

Орган по аккредитации должен являться юридическим лицом или быть созданной на базе организации — юридического лица; иметь необходимую инфраструктуру и экономические возможности для управления системой аккредитации; располагать беспристрастными специалистами необходимой квалификации, в том числе экспертами-аудиторами, прошедшими подготовку в качестве специалистов по аккредитации и т. д. Общие требования к органам по аккредитации изложены в международном документе [96].

Функции национального органа по аккредитации в Российской Федерации возложены на Федеральную службу по аккредитации (Росаккредитация) — юридически это отдел Министерства экономического развития Российской Федерации [97]. В то же время аккредитацию лабораторий успешно проводят и негосударственные органы по аккредитации, из которых наиболее авторитетным является ААЦ «Аналитика» — член ILAC (см. разд. 1.6.5). Аккредитация в национальном органе обычно требуется по отраслевым документам в областях обязательного государственного регулирования [16].

В конспективном изложении процедура аккредитации аналитических лабораторий выглядит следующим образом.



Лаборатория направляет в орган по аккредитации заявку на аккредитацию, сопровождаемую документами. В первую очередь это Положение о лаборатории, паспорт лаборатории и ее Руководство по качеству.

В документах отражаются:

- сведения о лаборатории общего характера (название и адрес, штат сотрудников и их квалификация, площадь занимаемых помещений и т. д.);
- анализируемые объекты и определяемые вещества (область аккредитации);
- используемые методики анализа и аппаратура;
- применяемые способы обеспечения и контроля качества (используемые стандартные образцы, способы внутрилабораторного контроля, участие в программах внешней оценки качества и т. д.);
- информация, подтверждающая независимость лаборатории (если лаборатория аккредитуется на независимость);
- другие сведения, нужные для предварительной оценки возможностей лаборатории в выполнении исследований в заявленной области аккредитации.

Орган по аккредитации проводит экспертизу представленной документации и в случае положительной ее оценки формирует комиссию, направляемую в лабораторию. Комиссия проверяет соответствие представленных сведений фактическому состоянию дел; проводит экспериментальную проверку технических возможностей лаборатории по составленной ею программе. Обычно предлагают в присутствии членов комиссии проанализировать какие-либо стандартные образцы. В случае их отсутствия используют другие приемы оценки характеристик анализа (см. разд. 3.4).

Если комиссия признает лабораторию технически компетентной и, иногда, независимой, орган по аккредитации принимает решение об ее аккредитации. Срок действия аккредитации ограничивается обычно пятью годами, после чего аккредитация проводится вновь. В течение этого срока предусматривается проведение периодического инспекционного контроля (аналогичный термин «подтверждение компетенции»), (фактически представляющих собой форму упрощенной аккредитации); при отрицательных результатах такого контроля аттестат аккредитации может быть отозван либо сокращена область аккредитации.

В последнее время все большее значение придают участию лабораторий в программах внешней оценки качества (proficiency testing или МСИ, см. разд. 4.6) — полученные оценки учитываются при аккредитации и при инспекционных проверках. Некоторые органы по аккредитации требуют такого участия в обязательном порядке.

Конечным результатом аккредитации является документ, подтверждающий компетентность и (иногда) независимость лаборатории, — *аттестат аккредитации*.

Финансирование аккредитации в большинстве случаев ведется за счет аккредитруемой лаборатории.

В контексте данной книги интересен ответ на вопрос: *является ли аккредитация способом повышения качества химического анализа?*

Имеется много свидетельств как того, что аккредитованные лаборатории работают зачастую хуже не аккредитованных, так и обратного. Первое, вероятно, связано с тем, что аккредитованная лаборатория в юридических спорах с заказчиками анализов «защищена» аттестатом аккредитации, в то время как не аккредитованная



вынуждена постоянно доказывать свое право на существование только высоким качеством исследований. В то же время лаборатории, готовясь к аккредитации, стараются улучшить внутренние процедуры обеспечения качества, часто приобретают оборудование и стандартные образцы, что способствует более высокому качеству исследований.

Другими словами, аккредитация лаборатории сама по себе не приводит к автоматическому повышению качества выполняемых ею исследований.

## ГЛАВА 4

# ОБЕСПЕЧЕНИЕ КАЧЕСТВА РЕЗУЛЬТАТОВ ХИМИЧЕСКОГО АНАЛИЗА

Качество количественного химического анализа можно определить как состояние анализа, при котором потребитель результатов («Заказчик») уверен в том, что погрешность (неопределенность) анализа соответствует установленным нормам или требованиям Заказчика с известным уровнем доверия, и она его устраивает. Из этого определения следует, что обеспечение качества затрагивает практически все аспекты деятельности лабораторий, выполняющих количественный химический анализ. Конечно, определяющую роль играет выбор методики КХА для решения той или иной задачи.

### 4.1. О выборе методик количественного химического анализа

Проблема выбора методики КХА встает как при организации новых лабораторий, так и в давно существующих лабораториях в ходе их модернизации (при замене устаревшего оборудования и методик анализа, при появлении новых задач и т. д.).

Химический анализ почти никогда не является самоцелью, а выполняется для решения какой-либо научной или практической задачи, которых существует огромное множество. Перечислим типы наиболее часто встречающихся задач:

- анализ сырья, полупродуктов и продуктов химических, металлургических и других производств;
- определение токсичных веществ в объектах окружающей среды;
- анализ проб минералов, горных пород, вод при геологоразведке и геохимических исследованиях;
- анализ биологических жидкостей (кровь, моча и т. д.) и тканей в медицинской практике и ветеринарии для диагностики заболеваний;
- анализ пищевых продуктов как с целью определения токсикантов, так и для оценки их питательной ценности;
- криминалистические исследования;
- определение содержания различных элементов и соединений в ходе научных исследований (при изучении свойств и структуры природных и вновь синтезируемых веществ) и т. д.

При рассмотрении любой из этих задач становится ясно, что ограничения на погрешности анализа, а также экономические требования к выбираемым методам формулируют не столько аналитики, сколько специалисты соответствующего профиля, исходя из своих требований и возможностей. Приведем несколько примеров.



При определении токсичных веществ в объектах окружающей среды и пищевых продуктах методика анализа должна обеспечивать измерение этих веществ при концентрациях, меньших ПДК (предельно допустимой концентрации) — как правило, 0,5 ПДК или еще более низких. Погрешность (неопределенность) анализа при концентрации определяемых веществ, близкой к ПДК, не должна быть слишком большой (обычно не более 25% отн.).

В случае анализа биологических жидкостей требования к погрешностям определяются необходимостью надежно дифференцировать здоровых и больных людей по результатам анализа. Чаще всего требуют, чтобы общая погрешность анализа не превышала 0,10–0,12 интервала, в котором лежит концентрация определяемого вещества для 90% всех людей в популяции. При этом важнейшую роль играет требование низкой стоимости оплачиваемых пациентами анализов. Кроме того, имеются и специфические ограничения: в клинко-диагностических лабораториях запрещено использование ядовитых и горючих веществ, сжатых газов и пр., не исключается невысокая квалификация лаборантов.

По отношению к криминалистическим исследованиям выдвигаются такие требования: возможность анализа микрообъектов и высокая надежность исследований, обеспечиваемая как методами анализа, так и высокой квалификацией исполнителей. При этом стоимость аппаратуры, как правило, не играет определяющей роли.

В рамках этой книги мы не рассматриваем детально вопрос о процедуре выбора методики количественного химического анализа, оптимальной для решения той или иной задачи. Заметим только, что это сложный процесс, в ходе которого учитывают:

- требования к погрешностям и пределу обнаружения методики КХА;
- стоимость анализа;
- квалификацию лаборантов, их способность выполнять те или иные процедуры;
- наличие в лаборатории того или иного оборудования (средств измерений, испытательного и вспомогательного оборудования);
- возможные специфические требования (например, упомянутый выше запрет на работу с ядовитыми, горючими и взрывоопасными веществами в клинко-диагностических лабораториях) и т. д.

Кроме того, часто тексты методик не соответствуют существующим нормативным документам, регламентирующим требования к аналитическим методикам и плохо написаны, что дополнительно затрудняет как выбор методики, так и ее документальное оформление в лаборатории. Поэтому в большинстве случаев процедура принятия решения о выборе методики не может быть полностью формализована и успех, в основном, зависит от знаний и опыта специалистов. Обычно методику выбирают из существующих аттестованных методик. Однако во многих случаях ее разрабатывают и аттестуют заново<sup>69</sup>. Интересующиеся данным вопросом могут обратиться к работам [22, 90, 98].

После того как методика выбрана из уже существующих либо разработана заново, ее реализуют в лаборатории («постановка методики»). При этом поставленную методику КХА исследуют, оценивая (полностью или частично) основные

<sup>69</sup> Об аттестации методик КХА см. разд. 3.3.

метрологические характеристики, перечисленные в разд. 3.2. Эту процедуру часто называют «валидацией методики в лаборатории». Ее проводят либо в специальных экспериментах, аналогичных рассмотренным в разд. 3.2, либо в ходе контроля стабильности (см. ниже, разд. 4.3).

## 4.2. Отбор пробы

Цель количественного химического анализа — измерить концентрацию тех или иных веществ в каком-либо объекте или его части. Часто используют термин «аналитический контроль».

*Аналитический контроль (объекта) — определение химического состава и, в отдельных случаях, структуры и свойств вещества и материала объекта аналитического контроля с последующим оцениванием соответствия объекта установленным требованиям при их наличии [99].*

При этом, как правило, непосредственно анализируется лишь малая часть исследуемого объекта, для чего достаточно от нескольких миллиграммов до нескольких граммов или максимум десятков граммов вещества. Рассматривая приведенный в п. 4.1 перечень наиболее часто встречающихся аналитических задач, легко понять, что в большинстве случаев масса исследуемых объектов гораздо больше, причем распределение определяемых веществ в них чаще всего неоднородно (фактически полностью однородными являются лишь истинные растворы). При этом чаще всего заказчика интересует информация о среднем содержании аналитов в объекте исследования. Иногда необходимы и сведения о распределении аналита в исследуемом объекте, например, распределение железа в рудном теле предполагаемого месторождения либо исследование распределения углерода по поверхности среза стального бруска.

### Примеры:

- требуется определить содержание железа в партии<sup>70</sup> железной руды (железнодорожный состав), имеющей массу несколько сот тонн и заведомо состоящую из кусков с различной концентрацией железа. Для анализа атомно-абсорбционным методом необходима навеска массой ~0,2 г;
- необходимо оценить среднее содержание глюкозы в сыворотке крови жителей г. С.-Петербурга в группах (стратах), отличающихся по возрасту и полу обследуемых;
- надо оценить содержание вредных для здоровья тяжелых металлов (свинец, мышьяк, кадмий и т. д.) в партии импортного картофеля массой 5000 тонн. Масса навески для анализа методом ICP-MS составляет 5 г сухого вещества;
- и т. д.

В некоторых случаях (при контроле изделий микроэлектронной техники, в криминалистике или при научных исследованиях) весь исследуемый объект представляет собой аналитическую пробу. Иногда аналитическую пробу не готовят, например при рентгенорадиометрическом анализе рудных материалов в условиях естественного их залегания, при непрерывном рентгеноспектральном анализе

<sup>70</sup> Под «партией [продукции]» понимают некоторое количество продукции, в том числе и сырья, произведенное одним поставщиком, характеризующееся одинаковой обработкой, имеющее одинаковое наименование и сорт, отправляемое в один адрес и сопровождаемое одним документом о качестве продукции.



шихты, перемещаемой на конвейерной ленте, и т. д. В остальных случаях от исследуемого объекта тем или иным способом отбирают некоторую часть (части), которую называют *пробой*.

*Проба — часть вещества (материала) объекта аналитического контроля, отобранная для анализа, и/или исследования его структуры, и/или определения свойств, отражающая его химический состав, и/или структуру, и/или свойства [99].*

*Отбор пробы — отделение части вещества (материала) объекта аналитического контроля с целью формирования пробы для последующего определения ее состава, структуры и/или свойств. Используют также термин «пробоотбор», что не рекомендуется.*

Часто применяется и близкий по смыслу термин «образец».

Проба (навеска, аликвота) должна иметь тот же состав, что и исследуемый объект. Например, проба воды из озера должна иметь тот же состав, что и вся вода в озере (или в какой-то его части), проба почвы, взятая для агрохимических целей, должна отражать содержание питательных веществ в почве обследуемого поля и т. д. Весьма типична ситуация, когда состав пробы существенно отличается от среднего состава исследуемого объекта и погрешность отбора пробы соизмерима с погрешностями количественного химического анализа (и даже превосходит их). Это отличие характеризуют погрешностью отбора пробы.

*Погрешность отбора пробы — отклонение значения величины, характеризующей состав, структуру, свойства пробы вещества (материала), от значения этой же величины, характеризующей состав, структуру, свойства объекта аналитического контроля в целом.* Погрешность отбора проб в основном определяется неоднородностью вещества или материала. Свой вклад вносят также погрешность результатов сопутствующих измерений, погрешность, обусловленную изменением состава, и/или структуры, и/или свойств пробы в процессе процедуры отбора пробы и др.

Часто говорят также о *неопределенности отбора пробы*, под которой понимают составляющую общей неопределенности, обусловленную процедурой отбора пробы вещества (материала) объекта аналитического контроля. Подробнее см. разд. 4.2.4 (см. также разд. 3.5.2 и сл.).

Пробу, отличающуюся в составе от состава исследуемого объекта можно пренебречь, называют *представительной*. Задача получения представительной пробы легко решается при исследовании однородных объектов — например, хорошо перемешанных жидкостей и газов. В остальных случаях процедура отбора пробы может оказаться весьма сложной и многоступенчатой.

Терминология отбора проб весьма развита. Различают следующие виды проб:

- в зависимости от **способа получения**: разовые, точечные (единичные, частные), мгновенные, суточные и т. п.
- в зависимости от **стадии первичной обработки**: исходные, промежуточные, объединенные, средние, сокращенные, лабораторные, аналитические и др.
- в зависимости от **назначения**: контрольные, рабочие, резервные, арбитражные и др.

Подробнее о видах проб см. ниже (разд. 4.2.1.2).

Часто говорят о *подготовке пробы* — совокупности процедур, проводимых с целью подготовки пробы вещества (материала) объекта аналитического контроля



к определению ее состава, и/или структуры, и/или свойств. Используя этот термин, необходимо четко определять, какие именно манипуляции с пробами имеются в виду. Дело в том, что под подготовкой пробы могут подразумеваться как операции, проводимые в лаборатории или вне ее до анализа — отбор проб и сопутствующие процедуры (высушивание, измельчение, перемешивание и т. д.) — так и обработка проб до измерения аналитического сигнала, являющаяся по сути частью собственно анализа (например, концентрирование и выделение аналита экстракцией или соосаждением).

Иногда весь комплекс операций на этапах отбора пробы и пробоподготовки называют *опробованием*.

Отбор пробы в большинстве случаев делится на два этапа:

- отбор пробы вне лаборатории, который заканчивается получением *лабораторной пробы* и передачей ее в лабораторию;
- отбор пробы в лаборатории, который состоит в отборе представительной части лабораторной пробы для собственно анализа (*аналитическая проба*).

#### 4.2.1. Отбор пробы вне лаборатории

Если исследуемый объект не доставляется в лабораторию целиком, отбор пробы начинается, естественно, вне лаборатории. Как правило, отбором пробы вне лаборатории и приготовлением лабораторной пробы (см. ниже) занимается особый персонал — это могут быть как сотрудники лаборатории, так и другие специалисты (например, геологи при геологических исследованиях или сотрудники особой группы в отделе главного технолога на заводе). При этом используется специальное оборудование. Все это должно выполняться в соответствии с утвержденной документацией (*«план и процедуры отбора пробы»*).

Независимо от того, кем именно проводится отбор и приготовление лабораторной пробы, крайне желателен контроль лаборатории за этим процессом, при котором фиксируются используемые для каждой пробы процедуры, выполнявшего их специалиста, условия окружающей среды (при необходимости), точки отбора проб и, если это целесообразно, статистические данные, лежащие в основе принятого плана отбора пробы. Это связано с тем, что погрешности отбора пробы в значительной степени определяют общую погрешность исследования (см. выше).

Конечно, возможны ситуации, когда лаборатория не информирована о деталях получения пробы. В этом случае в протоколе с результатами измерений рекомендуется указывать «отбор пробы выполнен Заказчиком».

Для достижения приемлемой погрешности отбора пробы используют различные способы.

##### 4.2.1.1. Основные пути отбора проб

При анализе веществ и материалов, представленных в больших количествах (объекты окружающей среды, руды и горные породы, партии товаров, большие совокупности объектов при биологических и медицинских исследованиях и т. д.), возможны два основных пути.

Один путь — отобрать некоторое число проб и проанализировать каждую пробу отдельно. Так поступают при популяционных биохимических исследованиях



в медицине. Из популяции (совокупности биологических объектов, в данном случае людей) делают выборку, отбирая случайным образом некоторое число обследуемых (например, выборка из жителей какого-либо города). Из крови каждого из них берут пробу и анализируют каждую из них отдельно. В этом случае среднее содержание определяемого вещества в популяции находят, усредняя результаты анализа индивидуальных проб. Достоинство такого подхода — возможность изучения распределения аналита в исследуемой популяции и анализ связи распределения аналита с другими показателями (так, например, была выявлена связь содержания холестерина в сыворотке крови с риском сердечно-сосудистых заболеваний).

Другой путь — усреднение большого числа проб, составление одной, т.н. генеральной пробы, последовательном отборе из нее меньших по массе представительных проб вплоть до получения лабораторной пробы, которую передают для анализа в лабораторию. Так поступают, например, при исследовании партии руды, перевозимой в вагонах, — из каждого вагона по определенному плану отбирают некоторое число проб, затем их поэтапно смешивают («объединяют») и отбирают из них представительные пробы меньшей массы («сокращают»), получая в конце концов лабораторную пробу. Промежуточный вариант часто применяют при геологоразведке — некоторое число проб, отобранных, например, на определенных участках обследуемого месторождения, объединяют и сокращают, получая некоторое число усредненных проб, анализируя которые получают информацию о месторождении.

В конечном счете выбор способа отбора пробы определяется задачами исследования (нужна или нет информация о распределении аналита в исследуемом объекте) и финансовыми соображениями — часто цена анализа одной пробы достаточно высока, и затраты на объединение и сокращение проб с лихвой окупаются.

#### 4.2.1.2. Виды проб

Рассмотрим, какие основные виды проб возможны в наиболее общем случае, когда процедура их отбора предусматривает последовательное объединение и сокращение.

Процесс начинается с составления *генеральной* (первичной, начальной, общей, суммарной, объединенной, исходной) пробы, которую получают объединением необходимого числа точечных (разовых, частных, единичных) проб (см. рис. 4.1). В геологии часто отбор пробы начинается с *валовых (объемных)* проб, каждая из которых представляет собой большой (обычно более 0,5 м<sup>3</sup>) объем материала, полученный при проходке горных выработок.

*Точечная* проба есть часть исследуемого объекта, которую отбирают за один прием (за одну операцию) из разных точек этого объекта в определенный момент времени в одном месте или на определенном уровне. Чаще всего точки отбора проб выбирают таким образом, чтобы их состав максимально отражал состав исследуемого объекта либо его части. Это наиболее трудоемкая и сложная часть процесса отбора пробы.

Если исследуются партии изделий, точечные пробы отбираются из отдельных изделий, обычно выбираемых из партии случайным образом.

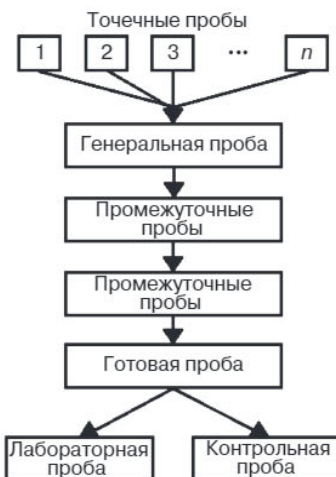


Рис. 4.1. Общая схема отбора проб

Поскольку масса отобранной генеральной пробы почти всегда бывает значительной (до 2–3% массы объекта исследований, что доходит до нескольких сотен килограммов), ее подвергают *разделке* (совокупность операций измельчения<sup>71</sup>, перемешивания и сокращения, см. ниже) по определенным правилам. При этом каждая операция может проводиться изолированно либо с использованием *технологических модулей* — устройств, состоящих из двух или нескольких синхронно работающих машин и позволяющих одновременно измельчать, перемешивать и сокращать пробу. Мы не рассматриваем конкретные способы разделки проб, подробно описанные в специализированной литературе (например, в [101–109]).

При использовании одного или нескольких циклов разделки получают промежуточные (или частичные) средние пробы, которые разделяют так же, как и генеральную пробу, в результате чего масса их последовательно уменьшается до тех пор, пока не будет получена готовая (средняя, сокращенная, товарная) проба.

Путем сокращения готовой пробы получают *лабораторную* (паспортную, сертификатную) пробу, предназначенную для проведения всех видов лабораторных испытаний и *контрольную* (арбитражную, архивную, дубликатную, резервную) пробу, которую хранят на случай проведения повторных, арбитражных или других контрольных испытаний.

*Лабораторная проба* — это конечная промежуточная проба, поступающая в лабораторию для анализа. По лабораторной пробе судят о составе исследуемого объекта (например, партии руды). Поэтому процедура ее отбора должна быть такова, чтобы состав лабораторной пробы не отличался бы достоверно от среднего состава как всех промежуточных и генеральной проб, так и всего объекта исследований. В зависимости от назначения масса лабораторной пробы различна. В среднем она колеблется от 0,5 до 2 кг (хотя бывают и пробы массой от 25 г до 30 кг). Готовую пробу обычно делят на две части, каждую из которых помещают в чистый, сухой и

<sup>71</sup> *Измельчение пробы* — механическое воздействие на материал пробы для уменьшения размера ее частиц.



герметично закрывающийся сосуд (металлические, стеклянные или пластмассовые банки). Одну из двух частей направляют в лабораторию для анализа (собственно лабораторная проба), а второй хранят на случай проверки (контрольная или арбитражная проба). Срок хранения контрольной (арбитражной) пробы зависит как от стабильности материала, так и от ситуации. Стабильные пробы хранят до тех пор, пока она еще может понадобиться в конфликтных ситуациях (как правило, в течение заранее оговоренного срока, составляющего обычно 1–6 месяцев).

По возможности пробы хранят в условиях, обеспечивающих их стабильность (исключающих воздействие света, влаги, кислорода и других факторов, которое может изменять пробы). Для стабилизации жидких проб часто добавляют консерванты (например, подкисляя их при определении тяжелых металлов или добавляя комплексообразователи, если они не мешают анализу). Биологические пробы (сыворотка крови, моча и т. д.) хранят в холодильнике или даже замораживают. Срок хранения проб оценивают экспериментально аналогично тому, как это делается для стандартных образцов состава (см. разд. 3.2.2).

Пробы газов обычно не хранят.

Опломбированные и маркированные сосуды с лабораторными пробами, сопровождаемые соответствующей документацией, передают в лабораторию.

#### 4.2.2. Отбор пробы в лаборатории

Подготовку проб к анализу проводят в соответствии с рабочим планом, составленным с учетом специфики пробы и метода анализа. Он должен содержать все необходимые сведения о пробах.

Обычно никаких трудностей не возникает, если проба целиком используется для анализа (например, определение металлов в куске ионообменной смолы, в котором адсорбированы ионы из некоторого объема воды). Проблема возникает в тех случаях, когда масса (объем) этой пробы, уже отобранной и, возможно, обработанной ранее, намного больше, чем требуется для анализа, а сама она существенно неоднородна.

В таких случаях разделкой переданной в лабораторию лабораторной пробы готовят *аналитическую пробу*, представляющую собой специально подготовленный материал (чаще всего порошок), из которого берут навески для выполнения основных и, при необходимости, контрольных измерений. Масса аналитической пробы должна быть достаточной для определения всех аналитов (в соответствии со спецификацией) — обычно от 5 до 20 г. Если это необходимо, одновременно готовят и *дубликат аналитической пробы* — материал того же состава, что аналитическая проба, который используют только для контрольных определений.

Собственно разделка лабораторной пробы для приготовления подготовленной для анализа аналитической пробы начинается после ее внешнего осмотра и регистрации в лабораторном журнале или компьютере<sup>72</sup> и для твердых проб может состоять из следующих этапов:

<sup>72</sup> При регистрации пробы фиксируются код, присвоенный ей заказчиком исследований, дата поступления, номер и дата сопроводительного документа, при необходимости состояние и фактическая масса пробы и т. д. Одновременно пробе присваивается внутренний порядковый номер, под которым она будет фигурировать в лаборатории.

- сушка;
- измельчение и перемешивание;
- сокращение.

**Сушка** необходима почти всегда как из-за того, что результаты анализа чаще всего приводятся в пересчете на сухое вещество, так и из-за трудностей с последующим измельчением влажного материала. Обычно пробы сушат в течение нескольких (3–5) часов при повышенной температуре. Для этого пробу рассыпают на противне тонким слоем и помещают в сушильный шкаф. Температура сушки может составлять от 50 до 120 °С в зависимости от устойчивости материала пробы.

Дальнейшая обработка нужна для получения настолько однородного материала, чтобы любая небольшая порция, взятая для собственно анализа, была идентична по составу любым другим порциям. Для этого лабораторную пробу измельчают до требуемого размера частиц и, если надо, сокращают для достижения минимально необходимой при данном измельчении массы, которая определяется расчетом (см. разд. 4.2.3) либо из специальных таблиц. Это и есть аналитическая проба. Конечно, ее должно быть достаточно для анализа.

Способов измельчения известно множество. Используют, например, небольшие шаровые либо шариковые мельницы. Однако чаще всего по старинке измельчение проводят в ступках (они бывают сделаны из твердых сплавов, фарфора, агата, яшмы и других твердых, износостойчивых материалов). При необходимости получившийся порошок последовательно просеивают через сита с постепенно уменьшающимися размерами отверстий, причем остаток на каждом сите дополнительно истирают и просеивают снова.

Степень измельчения пробы должна быть оптимальной. Плохо измельченная проба труднее переводится в раствор при обработке кислотами либо сплавлением, а чрезмерное измельчение может привести к изменению состава пробы вследствие износа материала ступки или мельницы, окисления компонентов пробы, адсорбции, потери или приобретения влаги и т. д. Обычно размер частиц измельченной пробы составляет 0,05–0,08 мм.

Способов измерения размеров частиц также много — ситовой анализ, счетчики Коултера, измерения с помощью микроскопов и т. д. Однако на практике прекрасно работает дедовский метод — измельчение считается достаточным, когда получается тонкий порошок (пудра), при втирании которого в поры пальцев не ощущаются крупные частицы. Конечно, для ядовитых веществ этот способ контроля лучше не использовать.

Измельченную лабораторную пробу сокращают методом квадратования или квартования. Для **квадратования** пробу высыпают на лист плотной бумаги, перемешивают и распределяют по бумаге тонким слоем, придавая веществу форму прямоугольника. Последний делят взаимно перпендикулярными линиями на 15–20 равных квадратов, из середины которых отбирают плоским шпателем в шахматном порядке порции на всю толщину слоя. Отобранные порции объединяют, перемешивают и вновь сокращают тем же методом, пока не достигнут необходимой массы пробы. При **квартовании** делают то же самое, но квадратов четыре.

Полученную в результате операций измельчения и сокращения аналитическую пробу помещают в герметичную упаковку (например, бюксы) и передают на анализ.



Дубликат аналитической пробы хранят некоторое время, установленное правилами предприятия. Перед взятием навески сосуд с аналитической пробой встряхивают для перемешивания пробы и отбирают навеску шпателем из 2–3 мест пробы на разной глубине, в том числе и со дна сосуда. Разделка жидких проб состоит только в хорошем перемешивании лабораторной пробы взбалтыванием и отборе аликвот.

Кроме этих общих приемов отбора и разделки проб для отдельных видов материалов применяются специфические приемы и правила. Особые сложности возникают при работе с нестабильными пробами.

#### 4.2.3. Сколько пробы нужно отбирать, чтобы она была представительной?

Это важнейший вопрос, на который надо ответить при планировании каждого этапа отбора пробы. Существует несколько подходов.

Полуэмпирическая *формула Ричардса – Чечотта*

$$w = kd^2, \quad (4.1)$$

где  $w$  — масса отбираемой пробы в кг;  $d$  — приблизительный диаметр наиболее крупных частиц в мм;  $k$  — коэффициент, зависящий от характера распределения полезного компонента в массе пробы, была выведена в начале XX века. Она широко применяется при отборе проб в геологии, однако пригодна для любых твердых объектов, состоящих из кусков (крупных частиц).

Величину  $d$  находят, измеряя размеры частиц. Коэффициент  $k$  связан с минимально необходимым числом частиц в отбираемой пробе. Чаще всего он принимается по аналогии и лежит в диапазоне:

- 0,05–0,1 при равномерном распределении аналита по исследуемому веществу;
- 0,1–0,2 при неравномерном распределении;
- 0,2–0,4 при весьма неравномерном распределении;
- 0,5–1,0 при крайне неравномерном распределении.

Если необходимо более точно определить коэффициент  $k$ , его находят экспериментально, оценивая погрешность отбора пробы при разных отбираемых массах аналогично тому, как это делается при исследовании стандартных образцов (см. разд. 3.2.3). Оптимальная масса — минимальная, еще обеспечивающая удовлетворительную погрешность отбора пробы.

Более точные результаты дает использование формулы *Демонда – Хальфердаля*:

$$w = Kd^a, \quad (4.2)$$

где  $w$  — масса отбираемой пробы, в кг;  $d$  — диаметр самых больших частиц;  $K$  — коэффициент, зависящий от характера исследуемого объекта,  $a$  — показатель степени, лежащий в диапазоне 1,5–2,6. Величины  $K$ ,  $d$  и  $a$  находят экспериментально в каждом случае, измеряя размеры частиц и отбирая пробы разной массы. Из-за трудоемкости определения коэффициентов  $K$  и  $a$  формула Демонда – Хальфердаля используется реже, чем формула Ричардса – Чечотта.

Неэмпирический подход к установлению необходимой для представительности массы отбираемой пробы был предложен в работе [110] и позже развит для



конкретных случаев (см., например, [111, 112]. Кратко изложим его основы, следуя [31, 110–112].

В простейшем случае предполагается, что исследуемый материал состоит из двух видов сферических частиц равной величины и плотности  $A$  и  $B$ . Обозначим процент частиц вида  $A$  как  $P_A$ , и процент частиц  $B$  как  $P_B = 1 - P_A$ . Предположим также, что определяется концентрация только одного аналита, его процентное содержание в частицах вида  $A$  и  $B$  равно  $C_A$  и  $C_B$  соответственно, и отбираемые пробы содержат  $n$  частиц каждая. Тогда средняя концентрация аналита в пробе составит

$$C = [nP_A C_A + n(1 - P_A)C_B] / P_A C_A + (1 - P_A)C_B, \quad (4.3)$$

а среднее стандартное отклонение содержания аналита в разных пробах, связанное с неоднородностью  $s_h$  —

$$s_h = |C_A - C_B| \left( \frac{P_A P_B}{n} \right)^{1/2}. \quad (4.4)$$

Для того чтобы погрешностью отбора пробы можно было пренебречь по сравнению с погрешностью анализа (см. ниже, формулу (4.3)), число частиц в пробе  $n$  должно удовлетворять условию

$$n \geq 3 \frac{P_A P_B (C_A - C_B)^2}{s_{\text{ан}}^2}. \quad (4.5)$$

Более точные результаты получаются с учетом плотности частиц:

$$n \geq 3 \left[ \frac{P_A P_B (C_A - C_B)^2}{s_{\text{ан}}^2} \right] \frac{d_1 d_2}{d}, \quad (4.6)$$

где  $d_1$  и  $d_2$  — плотности частиц типов  $A$  и  $B$ ,  $d$  — общая плотность пробы.

В большинстве случаев можно считать  $s_{\text{ан}} \approx \sigma_R$ .

Формулы (4.5) и (4.6) легко экстраполируются на случай нескольких видов частиц [111]. Однако это почти никогда не требуется — модель практически любого материала как смеси частиц двух видов достаточно точна. *Погрешности расчетов по формулам (4.6–4.9) обычно определяются недостаточно точным знанием величин  $P_A$ ,  $P_B$ ,  $C_A$  и  $C_B$  для реальных проб, что, в свою очередь, определяется большой дисперсией частиц по размерам и, особенно при малых размерах частиц, с их агрегацией.*

Формулы (4.3–4.6) выведены в предположении, что компоненты исследуемого объекта распределены случайным образом, то есть различные части его не связаны друг с другом. Чаше всего так и бывает, однако есть исключения.

Например, если отбираются образцы из разных точек емкости, в которой происходит смешивание двух материалов, то пробы, отобранные из находящихся рядом точек, ближе по составу, чем отобранные из точек, удаленных друг от друга. То же происходит при последовательном отборе проб из потока жидкости, вытекающей из резервуара с неоднородным содержанием. В подобных случаях пробы оказы-

ваются взаимно зависимыми — состав смежных проб отличается мало, а с ростом отдаленности различие увеличивается. Подобные пробы называют взаимно коррелированными, а объект исследования — имеющим внутреннюю корреляцию. Сила связи между разными частями объекта характеризуется корреляционным фактором.

Понятно, что в случае объектов с внутренней корреляцией число отбираемых проб может быть уменьшено по сравнению со случаем независимых проб. Способ оценки необходимой массы отбираемой пробы для таких случаев приведен в [113].

Заметим, что описанные в данном разделе закономерности можно использовать как для расчета необходимой массы отбираемой пробы, так и для оценки требуемой дисперсности образца при фиксированной массе пробы (например, в тех случаях, когда она задана методикой анализа).

#### 4.2.4. Оценка и учет погрешности (неопределенности) отбора пробы

С самой общей точки зрения дисперсию, характеризующую погрешность (неопределенность) исследования какого-либо объекта, можно представить в виде суммы

$$\xi_{\text{общ}}^2 = \xi_h^2 + \xi_{\text{ан}}^2, \quad (4.7)$$

где  $\xi_h^2$  — дисперсия, характеризующая погрешность отбора пробы,  $\xi_{\text{ан}}^2$  — дисперсия, характеризующая погрешность анализа. Как всегда, при работе с дисперсиями равенство (4.7) справедливо для дисперсий соответствующих генеральных совокупностей. Для оценок справедливо приближенное соотношение

$$s_{\text{общ}}^2 = s_h^2 + s_{\text{ан}}^2, \quad (4.8)$$

где  $s_{\text{общ}}^2$ ,  $s_h^2$  и  $s_{\text{ан}}^2$  — экспериментально получаемые оценки дисперсий, характеризующие общую погрешность, погрешность отбора пробы и погрешность анализа соответственно. В соответствии с принципом пренебрежения малыми погрешностями (см. разд. 2.6.1) погрешностью отбора пробы можно пренебречь в случае, когда

$$s_h \leq 1/3 s_{\text{ан}}. \quad (4.9)$$

Однако при исследовании объектов сложного состава добиться этого почти никогда не удастся (либо это слишком дорого), и погрешность отбора пробы зачастую вносит значительный, а иногда и определяющий вклад в общую погрешность исследования в соответствии с формулами (4.7, 4.8). Формулу (4.8) используют и при оценке общей неопределенности исследований (аналитического контроля) тех или иных объектов, включающих стадии отбора пробы и химического анализа (см. также разд. 3.5.2 и сл.).

Неопределенность отбора пробы можно оценить, составив бюджет неопределенности, аналогично тому, как это описано в разд. 3.5.2.3.

**Пример:** оценка неопределенности отбора аналитической пробы при определении концентрации платины в новом катализаторе.

Для оценки эффективности нового катализатора он испытывался в реальных условиях в течение нескольких дней, после чего было необходимо оценить в нем концентрацию платины. Для определения концентрации платины был выбран метод пламенного атомно-абсорбционного анализа. Масса исследуемого катализатора составляла 5 кг, масса навески

для выполнения анализа составляет 100 мг. Было принято решение выполнять отбор в два этапа — вначале создать промежуточную пробу массой ~100 г, объединив 8 точечных проб массой ~12,5 г каждая, после чего сократить ее, измельчив в лабораторной шаровой мельнице и отобрав лабораторную пробу массой 5 г. Дальнейшее измельчение и отбор аналитической пробы выполнялись в лаборатории с помощью агатовой ступки.

Для оценки стандартного отклонения неоднородности при отборе промежуточной пробы отобранные точечные пробы были проанализированы методом рентгенофлуоресцентного анализа (РФА). Рассчитанное стандартное отклонение  $s_{h,1}$  составило 0,10. Известно, что стандартное отклонение повторяемости (сходимости) РФА не превышает 0,02, то есть в данной ситуации в соответствии с принципом пренебрежения малыми погрешностями им можно пренебречь. Тогда стандартное отклонение, характеризующее промежуточную пробу, можно оценить как  $s_{h,2} = 0,10 / \sqrt{8} \approx 0,035$ .

Стандартное отклонение отбора лабораторной пробы было оценено исходя из размеров частиц промежуточной пробы и распределения платины между ними величиной  $s_{h,3} = 0,03$ .

Аналогичная оценка отбора аналитической пробы дала  $s_{h,4} = 0,02$ .

Таким образом, общее стандартное отклонение отбора пробы можно оценить как  $s_h = \sqrt{s_{h,2}^2 + s_{h,3}^2 + s_{h,4}^2} = \sqrt{0,035^2 + 0,03^2 + 0,02^2} \approx 0,05$ . Это и есть стандартная неопределенность отбора пробы  $u_{c,h}(X)$ .

Стандартное отклонение воспроизводимости определения платины методом пламенного атомно-абсорбционного анализа при содержании платины 0,1–5% (что соответствует возможным содержаниям в исследуемом катализаторе) оценивается как  $s_{R,ан} = 0,02$ . Такова же, очевидно, стандартная неопределенность анализа  $u_{c,ан}(X)$ . Общая стандартная неопределенность процедуры аналитического контроля составляет

$$u_c(X) = \sqrt{u_{c,h}^2 + u_{c,ан}^2} = \sqrt{0,05^2 + 0,02^2} \approx 0,054.$$

Заметим, что, как это очень часто бывает, в данном случае основной вклад в общую неопределенность аналитического контроля вносит стадия отбора пробы.

### 4.3. Статистически контролируемое состояние

Важнейшим требованием к измерениям является то, что они должны выполняться в *статистически контролируемом состоянии*. Фундаментальное понятие «статистически контролируемое измерение» восходит к работе [114], в которой было введено понятие *статистический контроль*. В настоящее время статистический контроль рассматривается как составная часть общего процесса управления качеством [22]. Можно утверждать, что вообще нельзя рассматривать метод исследования как измерительный процесс до тех пор, пока он не подвластен статистическому контролю [115].

Существует несколько определений статистически контролируемого измерения. Так, согласно [116], «процесс измерения тогда находится в статистически контролируемом состоянии, когда причины, вызывающие значительную изменчивость результатов измерений, исключены или скорректированы, причем это сделано таким образом, что серия, состоящая из конечного числа измерений  $N$ , происходящих из процесса, пригодна для того, чтобы:

- заранее прогнозировать границы варьирования числа измерений  $N$ ;
- обозначить доверительную вероятность для того, чтобы будущие измерения находились внутри этих границ».





Другое определение более формально [22]: «Если результаты измерений, полученные в процессе измерений, накоплены в рамках любых возможных, но фиксированных условий, и они ведут себя так, как будто бы их выбрали случайным образом из фиксированного распределения, которое имеет фиксированную область значений и фиксированную изменчивость, то процесс измерения находится в статистически контролируемом состоянии».

Фактически эти и другие определения статистически контролируемого измерения сводятся к следующему: процесс измерений должен быть таким, что если бы исследовался один и тот же объект, то все результаты измерений или любую их часть можно было бы рассматривать как случайную выборку из совокупности, распределение которой не должно меняться во времени. Доверять результатам химического анализа и вообще измерений, корректно обрабатывать их методами математической статистики можно только тогда, когда процесс измерений находится в статистически контролируемом состоянии и это подтверждено специальными методами. *Это общий принцип.*

С современных позиций все выглядит более конкретно. Предполагается, что если рассмотреть процесс измерений в данной лаборатории за длительное время, то он может быть рассмотрен как совокупность *серий*<sup>73</sup> измерений (в [27] использован оборот «результаты, полученные в условиях повторяемости (сходимости)», что для нас слишком длинно). Считается, что результаты, относящиеся к одной серии, получены в одинаковых условиях (время, калибровка, лаборант и прибор), а от серии к серии всегда изменяются одни и те же факторы (например, время и калибровка). В реальности это не всегда строго выполняется, однако такая модель достаточно проста и в то же время эффективна. Можно сказать следующее: мы предполагаем, что результаты последовательных измерений одного и того же неизменного и стабильного объекта в течение длительного времени описываются моделью

$$X_{j,i} = \alpha + \alpha_j + \varepsilon_{ji}, \quad (4.10)$$

где  $X_{j,i}$  — результат  $i$ -го измерения в  $j$ -й серии,  $j = 1, 2, \dots, J$  — номер серии измерений;  $i = 1, 2, \dots, I$  — номер измерения в конкретной  $j$ -й серии;  $\alpha$  — математическое ожидание измеряемой величины,  $\alpha_j$  — эффект серии и  $\varepsilon_{ji}$  — случайная погрешность. Предполагается нормальное распределение случайных эффектов —  $\alpha_j$  и  $\varepsilon_{ji}$  имеют нормальное распределение с нулевыми средними значениями и дисперсиями  $\sigma_{\alpha}^2$  и  $\sigma_{\varepsilon}^2$  соответственно:  $\alpha_j \in N(0, \sigma_{\alpha}^2)$  и  $\varepsilon_{ji} \in N(0, \sigma_{\varepsilon}^2)$ . Можно сказать, что эффект  $\alpha_j$  характеризует разброс результатов между сериями измерений, а эффект  $\varepsilon_{ji}$  — внутрисерийный разброс. Конечно, на самом деле измерения проводятся для разных образцов, но сказанное касается любого из них — если бы любой конкретный образец был бы всегда одинаков и исследовался бы многократно в течение длительного времени, результаты описывались бы моделью (4.10). Фактически, что это и есть определение статистически контролируемого состояния в современном понимании.

Коротко сформулировать вышесказанное можно следующим образом:

<sup>73</sup> Термин «серия» соответствует широко употребляемому английскому «run» и введен нами ранее, см. разд. 3.3.1.

*Статистически контролируемым называется такое состояние измерений (методики количественного химического анализа), при котором результаты многократных исследований одного и того же объекта (любого из стабильных, однородных и адекватных материалов в случае КХА) описываются моделью (4.10) с одними и теми же параметрами.*

В действительности среднее содержание определяемого вещества  $\alpha$  зачастую не постоянно, а изменяется во времени, т. е. его можно было бы записать как  $\alpha(t)$ . Дисперсии  $\sigma_{\alpha}^2$  и  $\sigma_{\epsilon}^2$  также могут изменяться во времени (обычно расти). Но поскольку целью контроля как раз и является выявление и предотвращение подобных изменений, приводящих к ухудшению метрологических характеристик анализа, то алгоритмы обеспечения подконтрольности измерений стараются строить таким образом, чтобы они выявляли как разовые грубые промахи, так и медленные изменения среднего (дрейфы) и дисперсий. Поэтому при построении таких алгоритмов рассматривают различные отклонения аналитического процесса от стационарного, описываемого моделью (4.10).

Конечно, обеспечение подконтрольности измерений — основная задача лаборатории, на решение которой направлена вся ее деятельность. То, что должно быть в хорошей лаборатории, описано в [73] — считается, что только выполнение всех требований этого документа обеспечивает статистически контролируемое состояние.

Важнейшим для обеспечения подконтрольности измерений является *внутрилабораторный контроль*. Основные способы такого контроля (см. разд. 4.5.1), вытекают из определения статистически контролируемого измерения и модели (4.10). В целом они сводятся к тому, что вначале достигается некоторое состояние измерений, а затем принимаются меры к тому, чтобы эти характеристики не ухудшались.

Понятно, что само состояние, которое должно быть стабильным, предполагается «хорошим», то есть устраивающим заказчиков исследований. Требования заказчика почти всегда — это требования нормативных документов на методики измерений, в которых обычно указываются нормативные величины — стандартные отклонения или пределы повторяемости (сходимости) и воспроизводимости для этих методик. Как это проверяется — отдельный вопрос, он рассмотрен нами ниже.

Практически все известные способы внутрилабораторного контроля базируются на одних и тех же представлениях о процессе измерений и статистической модели этого процесса [27].

Ниже мы постарались не просто дать набор рецептов для ведения внутрилабораторного контроля, но выделить наиболее последовательные и эффективные на сегодняшний день алгоритмы и методы контроля и комплексно и подробно рассмотреть их с учетом собственного опыта.

## 4.4. Проверка приемлемости

Процедура проверки приемлемости предназначена для выявления грубых промахов при анализе рутинных проб. Она в целом является частью внутрилабораторного контроля (см. ниже), однако вынесена нами в отдельный раздел, поскольку в отличие от других способов внутрилабораторного контроля:



- при проверке приемлемости оценивается качество исследования отдельных образцов, а не состояние методики в целом;
- процедура оценки приемлемости наряду с оценкой качества включает выполнение дополнительных измерений и получение окончательно приводимых результатов измерений.

Большинство методик КХА для каждой рутинной пробы предусматривают выполнение  $N = 2$ , реже  $N = 3 - 5$  измерений и обычные в статистике методы выявления грубых промахов, рассмотренные в разд. 2.4.2, либо неприменимы, либо малоэффективны. Действительно, при двух измерениях без дополнительной информации вообще нельзя судить, какое из них является грубым промахом, а при трех или четырех измерениях мощность стандартных методов крайне низкая — выявляются лишь сильно выпадающие результаты. В то же время в химическом анализе грубые промахи — хотя и сравнительно редкое, но неизбежное явление. Если не принять специальных мер, грубый промах автоматически приводит к неправильному результату анализа соответствующей рутинной пробы. *Собственно говоря, целью выполнения двух или более измерений для каждой рутинной пробы является получение возможности исключить влияние грубых промахов на окончательно приводимый результат анализа*<sup>74</sup>.

**Проверка приемлемости** — процедура оценки качества анализа каждой рутинной пробы. Ее применение позволяет гарантировать качество исследования рутинных проб, избегая грубых промахов. Проверку приемлемости результатов измерений проводят для каждой рутинной пробы, когда методикой предусматривается выполнение двух или более измерений для каждой пробы [27]. По результатам проверки приемлемости выбирают способ получения окончательно приводимого результата, который и вносится в протокол и передается Заказчику исследований.

При выполнении такой проверки предполагается, что:

- измерения для каждой рутинной пробы выполняются в одной серии;
- разброс результатов в пределах серии (эффект  $\varepsilon_{ji}$  в модели (4.10)) имеет нормальное распределение;
- стандартное отклонение повторяемости (сходимости)  $s_r$  не превышает приписанного значения  $\sigma_r$ , установленного в ходе аттестации (валидации) методики (см. разд. 3.4.3). То, что это справедливо, должно подтверждаться в ходе контроля стабильности (см. разд. 4.5.1).

Исходя из этого предполагается, что расхождение (размах) результатов измерений  $X_1, X_2, \dots, X_r, \dots, X_N$ , выполненных для каждой рутинной пробы  $R = (X_{\max} - X_{\min})$ , где  $X_{\max}$  и  $X_{\min}$  — наибольшее и наименьшее из величин  $X_r$ , описывается распределением размахов выборки из нормальной совокупности. Тогда 95% всех размахов не должно превышать величину  $CR_{0,95}(N)$ , где  $N$  — число измерений для одной рутинной пробы. Величины  $CR_{0,95}(N)$  находят, умножая стандартное отклонение повторяемости (сходимости) на коэффициенты  $f(N)$ , приведенные в табл. 4.1:

$$CR_{0,95}(N) = f(N)\sigma_r \quad (4.11)$$

Если  $n = 2$ , то  $CR_{0,95}(2)$  есть предел повторяемости (сходимости)  $r$ .

<sup>74</sup> Мы будем использовать терминологию [27]: результаты, получаемые в ходе измерений, будем называть просто «результаты», а результаты, получаемые в ходе их обработки и выдаваемые Заказчику, — «окончательно приводимые результаты».



Если размах не превышает  $CR_{0,95}(N)$ , то разброс результатов удовлетворителен и окончательно приводимый результат рассчитывается как среднее полученных результатов анализа. Если размах превышает  $CR_{0,95}(N)$ , это считается маловероятным и связанным с тем, что какой-то из результатов измерений  $X_{\max}$  или  $X_{\min}$  является грубым промахом<sup>75</sup>. В этом случае для получения окончательно приводимого результата используют специальные процедуры.

Таблица 4.1. Значения коэффициентов  $f(N)$  для расчета величин  $CR_{0,95}(N)$  формуле (21) [23]

$n$	$f(N)$	$n$	$f(N)$	$n$	$f(N)$
2	2,77	7	4,17	12	4,62
3	3,31	8	4,29	13	4,68
4	3,63	9	4,39	14	4,74
5	3,86	10	4,47	15	4,80
6	4,03	11	4,55	16	4,85

Рассмотрим, как все это делается в наиболее частом случае, когда для каждой рутинной пробы методикой предусмотрены 2 измерения ( $N = 2$ ).

Основной алгоритм показан на рис. 4.2. После получения двух результатов измерения для какой-либо рутинной пробы  $X_1$  и  $X_2$  рассчитывают размах (расхождение)  $|X_1 - X_2|$  и сравнивают его с пределом повторяемости (сходимости)  $r = CR_{0,95}(2) = 2,77\sigma_r$  – собственно проверка приемлемости.

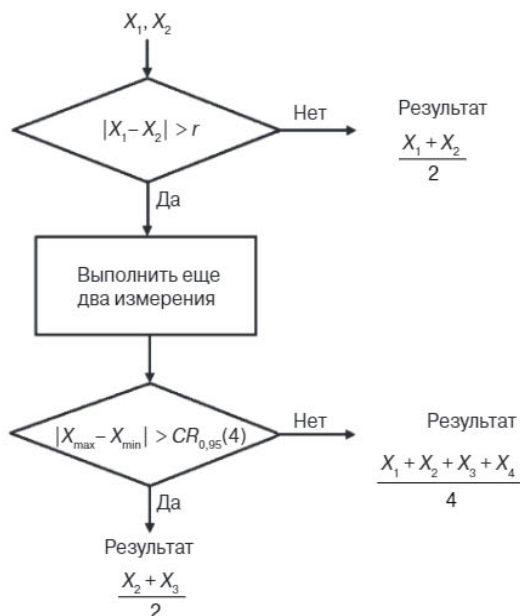


Рис. 4.2. Наиболее употребительный алгоритм проверки приемлемости и расчета окончательно приводимого результата для случая двух измерений каждой пробы ( $X_1$  и  $X_2$ )

<sup>75</sup> Предполагается, что для каждой пробы возможен только один грубый промах.

**Получение окончательно приводимого результата.** Если размах меньше  $r$ , то (окончательно приводимый) результат измерения для данной пробы рассчитывается как среднее из двух результатов измерений. В противном случае проводят два дополнительных измерения, после чего рассматривают все четыре результата как одну выборку. Сравнивают размах этой выборки с величиной  $CR_{0,95}(4)$  — критическим значением распределения размахов для 4 измерений и доверительной вероятности 0,95 (эта величина играет при четырех измерениях ту же роль, что и величина  $r$  при двух). Если рассчитанный размах меньше или равен  $CR_{0,95}(4)$ , то окончательно приводимый результат рассчитывают как среднее из четырех результатов, если больше — как медиану. В последнем случае грубый промах (предполагается, естественно, что им является одно из крайних значений в упорядоченной выборке — наибольшее или наименьшее) заведомо не участвует в расчете окончательно приводимого результата измерений.

В [27] приведены также другие алгоритмы проверки приемлемости и расчета окончательно приводимого результата.

Для случая, когда методикой предусмотрено два измерения для каждой рутинной пробы ( $N = 2$ ), возможны следующие варианты:

- использовать алгоритм, показанный на рис. 4.2;
- вместо двух выполнять одно дополнительное измерение (это делается при дорогих и длительных измерениях);
- выполнять одно, а при необходимости еще одно (для дорогих, но легко и быстро выполняемых измерений).

Если методика предусматривает  $N \geq 3$  измерений, можно:

- проводить дополнительно  $n$  измерений;
- выполнять дополнительно  $M$  (число  $M \leq N/2$ ) измерений;
- не выполнять дополнительных измерений.

Общий принцип остается одним и тем же: если размах не превышает критического значения распределения размахов для данного числа измерений, в качестве окончательно приводимого результата берется среднее значение всех результатов; в противном случае — получается необходимое количество результатов и рассчитывается медиана<sup>76</sup>. Описанные алгоритмы практически гарантируют получение надежных (не подверженных влиянию грубых промахов) результатов для всех рутинных проб.

## 4.5. Внутрिलाбораторный контроль качества

Под внутрिलाбораторным контролем качества измерений понимают совокупность действий, предпринимаемых аналитической лабораторией с целью подтверждения соответствия характеристик качества результатов анализа [материала] установленным требованиям [99].

Внутрिलाбораторный контроль качества необходим при выполнении химического анализа и других сложных измерений (под которыми мы понимаем

<sup>76</sup> Конечно, при нормальной работе случаи расчета окончательно приводимого результата как медианы достаточно редки (см. также разд. 4.5.2).

измерения, предполагающие, как и большинство методов КХА, сложные, часто многостадийные методики с большой долей ручного труда). Особенности этих измерений делают их потенциально нестабильными, подверженными ошибкам разного рода, и гарантировать их подконтрольность (см. ниже) можно только проводя специальные мероприятия. Конечно, при измерениях вообще и при количественном химическом анализе в частности почти всегда контролируются условия измерений. Однако, как показывает практика, этого недостаточно. *Внутрилабораторный контроль совершенно необходим для лабораторий, систематически выполняющих однотипные рутинные анализы (заводские лаборатории, лаборатории, осуществляющие анализ объектов окружающей среды, клинко-диагностические лаборатории и т. д.).* Вместе с тем такой контроль также крайне желателен и во многих научных исследованиях, в частности для обеспечения лучшей сопоставимости результатов, получаемых в разное время.

Ниже рассмотрение алгоритмов и процедур внутрилабораторного контроля ведется для методик количественного химического анализа. Однако это не означает, что они непригодны для «классических» измерений, рассмотренных нами в гл. 1 (то есть измерений величин, для которых построена система передачи размера от первичного эталона к образцовым и рабочим средствам измерений, таких как масса, длина, напряжение и т. д., не требующих сложной методики измерений). Отечественная метрология средств измерений основана на системе поверок (см. гл. 1) и не предполагает внутрилабораторного контроля. Считается, что метрологические характеристики средств измерений контролируются в ходе периодических поверок и гарантированно сохраняются в течение межповерочного интервала [3]. На практике это не всегда так, особенно для технически сложных устройств. Контроль стабильности мог бы гарантировать статистическую подконтрольность средств «классических» измерений и увеличить межповерочные интервалы (либо поверять средства измерений только при выходе из подконтрольного состояния). Именно поэтому в основополагающем документе [27], описывающем как терминологию и методы оценки погрешностей, так и, в шестой его части, процедуры внутрилабораторного контроля, не делается различий между разными видами измерений (хотя все приведенные примеры относятся к химическому анализу). Действительно, ничто не мешает контролировать стабильность весов, периодически взвешивая одну и ту же гирю.

*Ниже, если это не оговорено особо, мы будем говорить о контроле качества количественного химического анализа.*

Можно выделить следующие основные способы внутрилабораторного контроля качества:

- *контроль стабильности измерений с помощью контрольных карт;*
- *оценку метрологических характеристик измерений за тот или иной период;*
- *проверку приемлемости результатов измерений (см. выше, разд. 4.4)<sup>77</sup>.*

<sup>77</sup> Строго говоря, при проверке приемлемости контроль повторяемости (сходимости) по размахам результатов измерений рутинных проб — является лишь частью процедуры. Остальное — это расчет окончательно приводимого результата измерений (см. разд. 4.4).



#### 4.5.1. Контроль стабильности

Основным инструментом внутрिलाбораторного контроля, рекомендованным основными международными документами, является *контроль стабильности* с помощью контрольных карт. Он основан на анализе результатов регулярных измерений какого-либо стабильного показателя, характеризующего результаты измерений, причем измерения рекомендуется выполнять в каждой серии. Лучше всего, когда таким показателем является собственно измеряемый показатель — в химическом анализе это концентрация определяемого вещества (веществ) в *контрольном материале* (см. разд. 3.1). Возможен также контроль стабильности в отсутствие контрольного материала, что хуже. При других измерениях для контроля стабильности используется любой объект с постоянными свойствами.

*Основными достоинствами контроля стабильности являются:*

1. *Рассмотрение процесса измерений в динамике. При принятии решения о необходимости принятия мер по улучшению качества измерений принимается с учетом, в том числе, того, как происходили измерения в недавнем прошлом;*

2. *Ведение контроля с опорой на метрологические характеристики, полученные в конкретной лаборатории. С учетом того, что они, как правило, лучше приписанных (приведенных в тексте методики), лаборатория получает некоторый запас качества.*

Помимо контроля по определяемому показателю в КХА могут использоваться и другие: аналитический сигнал образца сравнения, используемого при градуировке; величины, по которым контролируется качество отдельных стадий анализа (например, масса или объем отобранных на одной из стадий химического анализа аликвот) и т. д. Нужно лишь, чтобы эти характеристики были стабильными в подконтрольном состоянии и отражали бы качество измерений. Мы будем рассматривать варианты контроля стабильности по результатам измерения определяемого данным методом показателя, однако все сказанное применимо и к контролю по другим величинам.

Изложенные ниже варианты контроля стабильности предназначены для оценки разных компонент погрешности, описываемых моделью (4.10). Так, контрольные карты размахов контролируют неизменность дисперсии  $\sigma_{\text{в}}^2$ , карты текущих расхождений — в основном дисперсии  $\sigma_{\text{а}}^2$ , а карты средних значений и кумулятивных сумм — главным образом погрешность, определяемую суммой эффектов  $\alpha$  и  $\alpha_j$ .

Перейдем к основному варианту контроля стабильности методик КХА, основанному на регулярном анализе *контрольного материала (материалов)*. При этом рассмотрим, как именно организуется такой контроль в различных ситуациях, а также алгоритмы оценки состояния измерений перед началом собственно контроля. Следует иметь в виду, что все сказанное может быть приложено к любым измерениям с небольшими оговорками — например, при контроле стабильности измерительных приборов в качестве контрольного материала может выступать мера.

Более слабые способы, применяемые в отсутствие контрольного материала, рассмотрены в разд. 4.5.1.4.

Общая схема контроля стабильности с использованием контрольных материалов следующая. Для каждого вида химического анализа (измерений):

1. Выбирают контрольные материалы.
2. Выбирают подлежащие контролю определяемые показатели.
3. Определяют, что является серией (аналитической серией, серией измерений).
4. Составляют схему (план) контроля, в которой задаются:
  - природа и количество контрольных материалов;
  - контролируемые показатели;
  - число измерений каждого показателя в каждом контрольном материале в одной аналитической серии.
5. В пределах некоторого количества аналитических серий проводят анализ (исследование) контрольного материала (материалов) для оценки реальных статистических характеристик аналитической методики. Мы будем называть эту стадию «период оценки характеристик».
6. По окончании периода оценки характеристик в случае соответствия погрешностей анализа предъявляемым к нему требованиям переходят к собственно контролю. При этом, насколько это возможно, отдельно контролируют точность в целом и повторяемость (сходимость) результатов анализа<sup>78</sup>.

Рассмотрим требования, предъявляемые к каждому из этих этапов, и наиболее эффективные алгоритмы контроля. Здесь мы будем рассматривать способы контроля стабильности, в которых результаты определения каждого показателя в каждом контрольном материале оцениваются отдельно<sup>79</sup>.

#### 4.5.1.1. Планирование процедур контроля стабильности

План эксперимента при внутрилабораторном контроле определяется как чисто метрологическими, так и экономическими факторами. *Мы будем полагать, что контроль стабильности методик КХА ведется по концентрации определяемых веществ в контрольном материале.* При контроле по другим показателям подходы в основном аналогичны.

**Выбор контрольных материалов.** Общие требования к контрольным материалам — стабильность их в течение срока применения, однородность и адекватность рутинным пробам. Наилучшими с точки зрения адекватности являются материалы одной природы с рутинными пробами (при анализе какого-либо сплава — это сплав близкого состава, при анализе сыворотки крови человека — стабилизированная лиофилизацией или замораживанием сыворотка крови человека и т. д.). Таким образом, общие требования к контрольным материалам близки к требованиям, предъявляемым к материалам, подготовленным для аттестации в качестве стандартных образцов состава [56].

<sup>78</sup> Иногда считают, что, если методикой предусмотрено одно измерение для каждой рутинной пробы, то такой характеристики, как повторяемость (сходимость), нет. Это явное недоразумение — характеристика методики, описывающая разброс результатов внутри серии, существует объективно, независимо от числа измерений.

<sup>79</sup> Один контрольный материал может использоваться для контроля качества нескольких показателей. Поскольку результаты контроля по каждому показателю оцениваются отдельно (совместный контроль нескольких показателей описан в разд. 4.5.1.3, «о многомерных методах контроля стабильности с использованием контрольных карт»), все нижесказанное справедливо для определения каждого показателя.



С позиций обеспечения надежности контроля количество контрольных материалов с разной концентрацией определяемого вещества должно быть по возможности велико, причем эти концентрации должны перекрывать всю область, в которой «работает» контролируемая методика химического анализа. Однако в реальности лаборатории ограничены экономическими соображениями и ассортиментом материалов.

Строгое решение задачи об оптимальном выборе числа контрольных материалов и концентраций определяемых веществ в них сопряжено со специальными дорогостоящими исследованиями. Цель таких исследований (которые могли бы проводиться, но обычно не делаются из-за их дороговизны) можно сформулировать так: надо найти такой минимальный набор материалов, чтобы стабильность, продемонстрированная при их анализе, гарантировала бы стабильность во всем диапазоне определяемых по данной методике концентраций.

*В итоге контроль стабильности чаще всего ведут с использованием одного, реже двух контрольных материалов.*

На практике контрольные материалы выбирают исходя из априорных знаний о применяемом методе (методике) анализа, а также с учетом важности того или иного диапазона концентраций для конкретных областей аналитических работ. При этом если концентрацию невозможно варьировать произвольно, контрольные материалы приходится выбирать из имеющихся в наличии. Так, при экологических исследованиях используют либо один материал с концентрацией вблизи ПДК, либо два контрольных материала с концентрациями вблизи и существенно выше ПДК; такой же выбор чаще всего оптимален в случае сертификации пищевых продуктов. Если весь диапазон концентраций одинаково важен, используют либо один материал с концентрацией вблизи середины диапазона определяемых концентраций, либо два или даже три контрольных материала (в последнем случае с содержанием аналита, находящимся на краях и в середине диапазона определяемых концентраций (медицина)).

Для контроля стабильности не обязательны (хотя и желательны) материалы с точно известным содержанием аналитов. Хотя очень часто для этих целей используют аттестованные материалы (ГСО, стандартные образцы предприятий, аттестованные смеси и т. д.), в случае их отсутствия либо из экономии берут не аттестованные — лишь бы они были стабильны, однородны и адекватны объектам анализа.

#### **Примеры:**

- контроль стабильности определения концентрации примесей в стали методом искрового эмиссионного анализа с использованием СО углеродистой стали ГСО 666-81П;
- контроль стабильности определения концентрации креатина в моче с помощью слитой мочи, стабилизированной антимикробными добавками.

В сфере клинической лабораторной диагностики, где контроль стабильности является обязательным, фирмы — производители стандартных образцов в больших количествах выпускают не аттестованные контрольные материалы — цена их в разы ниже цены аналогичных стандартных образцов.

Отдельно надо рассмотреть ситуацию, когда не удастся приобрести или изготовить контрольный материал, адекватный исследуемым пробам, то есть невозможно



вести контроль стабильности методики в целом. При этом, конечно, остается возможность вести контроль без контрольного материала (такие варианты рассмотрены в разд. 4.5.1.4). Однако во многих случаях в подобных ситуациях все-таки возможен частичный контроль стабильности — с материалами, имитирующими вещества, получающиеся после каких-либо этапов выполнения методики. Например, при определении концентрации тяжелых металлов в пылевых частицах в воздухе рабочей зоны изготовление контрольного материала, адекватного исследуемому объекту, практически невозможно. При этом легко изготовить стабильный раствор, похожий на те, которые получаются после растворения пыли из аспиратора, и вести по нему контроль стабильности. Конечно, в таком случае не будут контролироваться стадии отбора пробы и перевода ее в раствор, однако стадия собственно анализа будет под контролем.

**Определение перечня контролируемых показателей.** Конечно, если по данной методике определяется лишь один показатель, вопрос о выборе не стоит. Однако в последние годы появляется все больше методов, в которых в ходе одного анализа определяется несколько (иногда до нескольких десятков) показателей: газовая хроматография, эмиссионный спектральный анализ, масс-спектрометрия и т. д. Конечно, можно контролировать стабильность по всем показателям, однако из соображений экономии их число во многих случаях можно ограничить. Делается это на основе знаний о природе метода и собственного опыта. Так, в газовой хроматографии факторы, влияющие на стабильность, практически одинаковы для разных определяемых веществ: воспроизводимость ввода пробы, неизменность состояния колонки и детектора и условий их работы и т. д. В подобных случаях можно выбрать один или два определяемых показателя (обычно самых «капризных») и контролировать по ним стабильность измерений, полагая, что если для выбранных веществ все стабильно, то и для остальных это тоже так.

**Пример.** На автоанализаторе ф. Термо (это газовый хроматограф с предварительным сожжением пробы в динамической вспышке) в различных объектах одновременно определяется концентрация четырех элементов: C, H, N, S. В лаборатории автора контроль стабильности этих исследований ведется только по результатам определения водорода в ацетанилиде. Из опыта известно, что концентрация водорода — самый нестабильный показатель из перечня определяемых элементов, а чистый ацетанилид — превосходный контрольный материал с точно известным содержанием водорода.

**Определение серии измерений (аналитической серии).** Согласно определению серии (см. разд. 3.3.1), это — совокупность образцов, исследуемых одним и тем же методом, в одной и той же лаборатории, при одной и той же калибровке, одним и тем же оператором, с использованием одного и того же оборудования, в пределах короткого промежутка времени (то есть в условиях повторяемости (сходимости) согласно [27]). В большинстве случаев в реальной лаборатории определение серии не вызывает трудностей — это некоторое количество проб, исследуемых (анализируемых) за один раз. Например, при определении концентрации тяжелых металлов в водных пробах методом эмиссионной спектроскопии с индукционно-связанной плазмой (ICP-ES) лаборант включает прибор, выполняет градуировку и «прогоняет» пробы, поступившие в лабораторию, одна за другой — это серия. Число рутинных образцов в серии («исследованных в условиях повторяемости (сходимости)»

ограничено сверху как реальными потребностями, так и неизменностью условий измерений.

Смысл понятия «серия измерений» в том, что мы должны быть уверены: все факторы (время, градуировка, лаборант и прибор) одинаковы для всех проб в серии. Пробы другой серии исследуются в других условиях; именно поэтому *контроль стабильности ведут в каждой серии* — иначе нельзя гарантировать качество результатов (а вдруг серия, в которой не было контрольных проб, была «плохой»?).

Сложности с определением, что такое серия, возникают, когда измерения ведутся без разделения проб на группы. Реальный пример: в заводской лаборатории химического производства необходимо постоянно и оперативно контролировать вязкость полупродуктов. Пробу отбирают, приносят в лабораторию и свободный в этот момент лаборант сразу ее исследует на любом свободном вискозиметре. В подобных ситуациях ищут компромисс: не следуя строго определению серии, понимают под ней пробы, проанализированные в одних условиях. В нашем примере это, скорее всего, все пробы, исследованные в течение одной смены — можно условиться, что в одну смену условия одинаковы, а в другую смену они будут другими (хотя возможны и варианты — например, считать, что серия — это пробы, полученные одним лаборантом в одну смену, что сложнее).

Другая непростая ситуация — непрерывно работающий автоматический прибор, выполняющий периодические (например, раз в 10 минут) измерения без участия человека.

**Пример** — круглосуточный контроль концентрации в химическом реакторе автоматическим хроматографом. Размер серии в этом случае определяют исходя из соотношения скорости возможных изменений статистических характеристик анализа (дрейфов) и общей погрешности исследования таким образом, чтобы этими изменениями можно было бы заведомо пренебречь. Другими словами, это — типичный период, в течение которого вероятность выхода процесса измерений из-под контроля достаточно мала.

Если есть возможность планировать размер серии, необходимо принимать во внимание и экономические соображения. Так, в случае «длинных» серий (20–30 и более рутинных проб) доля затрат на контрольные материалы и их анализ значительно меньше, чем при «коротких».

**Выбор числа измерений для контрольного материала в одной серии.** Когда контрольный материал (материалы) для контроля данной методики измерений выбран и определены показатели, по которым будет проводиться контроль, остается определить число измерений, которые будут выполняться для каждого контрольного материала в каждой серии, и расположение контрольных проб в сериях. *В дальнейшем число измерений для данного контрольного материала и определяемого показателя нельзя изменить без нового набора статистики или, по крайней мере, нового расчета статистических характеристик (при уменьшении числа измерений).*

Во многих случаях число измерений для контрольного материала выбирают таким же, как и при анализе рутинных проб. На первый взгляд это кажется естественным (и даже включено в нормативные документы [117]). На самом деле ничто не мешает (и даже желательно) выполнять больше измерений для контрольного материала, чем для рутинных проб. Это наиболее очевидно, когда при поточном рутинном анализе из соображений экономии проводят только одно измерение для



каждой пробы (анализ почв, клинический анализ, геологические исследования и т. д.). В этом случае на результат измерений влияет как межсерийный разброс — промежуточная прецизионность, так и случайный разброс внутри аналитической серии — повторяемость (сходимость), причем эти факторы имеют разную природу. Поскольку целью контроля стабильности является не только выявление неудовлетворительных результатов, но и устранение причин возникшего сбоя, лучше контролировать оба компонента погрешности — и повторяемость (сходимость), и промежуточную прецизионность — для чего надо выполнять два или более измерения.

Дополнительным аргументом, заставляющим стремиться к возможно большому числу параллельных измерений, является тот факт, что согласно центральной предельной теореме [118] даже при существенном отклонении распределения результатов единичных измерений от нормального распределение среднего из нескольких результатов приближается к нормальному (при этом чем больше результатов усредняется, тем ближе распределение средних к нормальному). В реальности при анализе вещества группировка с усреднением приводит к нормальным распределениям средних значений уже при числе измерений в группе 4–5, а часто и меньше [94, 118]. А практически все методы внутрилабораторного контроля основаны на предположении о нормальном распределении результатов измерений, см. модель (4.10).

С другой стороны на организацию рутинных измерений и контроль их качества в огромной степени влияют экономические факторы. Исследование контрольных материалов может вносить существенный вклад в стоимость измерения для рутинных проб, и этот вклад прямо зависит от числа измерений для контрольных материалов. Минимальное число измерений для одного контрольного материала, позволяющее контролировать повторяемость (сходимость) по контрольной карте размахов — два. *Именно поэтому чаще всего при контроле стабильности проводят именно два измерения для каждого контрольного материала.* То же мы видим в [27], где во всех примерах фигурирует по два контрольных измерения.

Важную роль играет расстановка контрольных проб в серии. Она выбирается из того соображения, что результаты контроля должны отражать качество исследования рутинных проб. Поскольку случайный характер погрешностей в пределах серии не гарантирован (возможны дрейфы и другие систематические эффекты, которые также надо выявлять), рекомендуется случайная расстановка. Идеально использовать таблицы случайных чисел. Часто применяемое, особенно при работе с автоматическими пробоотборниками, расположение контрольных проб на фиксированных позициях может привести к некорректным оценкам. В нашей практике были случаи, когда из-за дефекта автоматического пробоотборника результаты измерений в некоторых позициях были неправильными (а в остальных — нормальными). При случайной расстановке контрольных проб такая погрешность рано или поздно (когда контрольная проба окажется в дефектной позиции) будет выявлена. Если же контрольные пробы всегда располагаются на одних и тех же позициях, этого может не произойти никогда.

**Некоторые организационные аспекты.** В целом можно выделить три составляющих внутрилабораторного контроля: персонал, который собственно выполняет



контроль; контрольные материалы (если контроль ведется с их применением, как чаще всего бывает); обработка и использование его результатов — ведение контрольных карт, выполнение расчетов и т. д. Часто создается иллюзия, что наиболее сложная компонента — последняя, несколько проще решить вопрос с контрольными материалами, а люди, конечно, будут делать все так, как надо. На самом деле все обстоит с точностью до наоборот: проще всего наладить расчеты и ведение контрольных карт, во многих случаях приходится помучиться с выбором контрольных материалов, а самое сложное — добиться, чтобы персонал честно и корректно выполнял нужные процедуры.

Дело в том, что все, что делается в рамках внутрिलाбораторного контроля (исследование контрольных материалов, обработка их результатов и т. д.), лаборанту<sup>80</sup> часто кажется дополнительной и ненужной работой. Соответственно, лаборанты часто делают эту работу спустя рукава, а иногда даже не делают ее вовсе, придумывая результаты контроля. Чтобы этого не происходило, необходимо:

- максимально удобно организовать контроль на рабочем месте;
- добиться понимания важности этих процедур всеми сотрудниками лаборатории.

Выполнение этих условий и определяет успех мероприятий по внутрिलाбораторному контролю.

Удобство выполнения контроля — естественное эргономическое требование. Конечно, если контрольный материал просто стоит на полке над рабочим столом, никаких проблем не возникает. Однако контрольные материалы часто отличаются от рутинных проб и требуют специальных условий хранения (например, в холодильнике). Часто при их исследовании необходимы специальные подготовительные процедуры, не нужные для рутинных проб (вскрытие ампул, размораживание, разбавление концентрата и т. д.), которые должны доставлять лаборанту минимум неудобств. Так, холодильник с контрольными пробами должен находиться в том же помещении, в котором происходят измерения, а не на другом этаже. То же касается и компьютерной обработки результатов измерений — компьютер либо терминал должен находиться рядом с рабочим местом лаборанта (благо сейчас это недорого). Все это кажется очевидным, но часто не делается — сказывается некоторое пренебрежение контрольными процедурами. Между тем, если выполнение контроля требует серьезных дополнительных усилий, его могут просто не делать.

Возможно, еще более важным является психологический аспект. Прямым путем к добросовестному и качественному выполнению процедур внутрिलाбораторного контроля является убеждение лаборантов и вообще всех сотрудников лаборатории в том, что их благосостояние в настоящем и будущем зависит от качества выполняемых исследований, а это качество невозможно гарантировать без внутрिलाбораторного контроля. На самом деле это действительно так. Для коммерческих лабораторий, работающих на свободном рынке, доходы (а значит, зарплата сотрудников и условия их работы) напрямую зависят от качества измерений — это очевидно. Однако и для заводских лабораторий выдача некачественных результатов чревата

<sup>80</sup> Под «лаборантом» мы понимаем любого сотрудника, выполняющего рутинные измерения.

ухудшением отношения со стороны руководства и, как следствие, уменьшением стимулирующих выплат и закупок оборудования и, возможно, кадровыми решениями. Даже один ошибочный результат измерения (если ошибка будет выявлена, а это происходит не так уж редко) потребители результатов — технологи, экологи и т. д. — будут вспоминать годами!

Эта мысль должна быть доведена до всех сотрудников лаборатории, а особенно — до лаборантов. Сделать это непросто, требуется индивидуальный подход. Однако можно дать общий совет — ни в коем случае нельзя ругать лаборанта и (тем более) наказывать его материально в случае неудовлетворительных результатов контроля: это будет тем, что психологи называют «отрицательным стимулированием». После «разноса» лаборант будет делать все возможное, чтобы он не повторился. А поскольку сбои в процессе измерений неизбежны и являются нормальным явлением, избежать их полностью можно только одним способом — подделывать результаты в случае выхода процесса из-под контроля. Конечно, картина при этом будет благостная, но вряд ли это улучшит реальное качество измерений.

Следует иметь в виду, что на самом деле низкое качество — следствие неудовлетворительного руководства; система руководства есть причина большинства проблем в этой области. Ключ к хорошему качеству — уважение, признание людей [22]. Поэтому нормальная реакция на выявленный выход процесса измерений из-под контроля — не претензии, угрозы и репрессии, а изъявление удовлетворения по типу «как хорошо, что мы вовремя заметили этот сбой и неправильный результат не пойдет заказчику. Теперь нам надо разобраться, из-за чего это произошло» и т. д. Такой образ действий создает максимально комфортную атмосферу в лаборатории.

В настоящее время во многих отечественных лабораториях психологические проблемы успешно решаются, и внутрилабораторный контроль проводится реально и в полном объеме.

**«Юридическая» проблема.** Одна из сложностей, возникающих при внедрении современных процедур внутрилабораторного контроля в нашей стране, состоит в том, что в методиках в соответствии с [6] прописаны процедуры контроля качества, и для большинства методик они безнадежно устарели. Общее число методик измерений оценивается в 30–35 тысяч и многие были разработаны еще в 70-е — 80-е годы прошлого века. Правда, приказ о внедрении документа [27] содержал требование приведения в соответствие с ним всех нормативных документов (в том числе и документов на методики), однако оно не было выполнено. Конечно, было бы правильно внести изменения в методики и узаконить их. Однако ясно, что этого не будет — объем работы огромен, да и многих организаций-разработчиков методик уже не существует. Получается, что во многих случаях использование современных методов контроля (например, контроль стабильности, о котором при разработке старых методик не имели понятия) формально противоречит тексту методик.

Вне сферы государственного контроля особой проблемы нет — легко внести соответствующие изменения в методику, узаконить их в рамках предприятия и согласовать с заказчиками исследований, тем более что в большинстве случаев контроль стабильности «поглощает» другие методы контроля. Согласно Закону о техническом регулировании [16] обязательными являются только технические регламенты — вы-



полнение требований всех остальных документов, включая методики измерений (если они не вошли в какой-либо регламент) является добровольным и во главу угла ставятся требования Заказчика. Поэтому лаборатории могут самостоятельно изменить методику в части процедур контроля качества измерений, согласовав эти изменения с заказчиками исследований и, при необходимости, утвердив эти изменения в рамках своего предприятия (стандарт предприятия). Конечно, все это должно отражаться в Руководстве по качеству лаборатории.

Сложнее обстоит дело с работами, лежащими в сфере государственного регулирования (см. разд. 1.6). Статьей 15 Закона [2] установлен государственный надзор за соблюдением аттестованных методик измерений, относящихся к сфере государственного регулирования обеспечения единства измерений. Формально, а это стандартный подход надзорных органов, изменение процедур контроля может быть расценено, как нарушение требований методики<sup>81</sup>. Понятно, что при аккредитации орган по аккредитации может закрыть на это глаза или даже указать, что процедуры контроля, записанные в методике, являются недостаточными, но неприятности со стороны надзорных органов вполне возможны. Остается надеяться на здравый смысл проверяющих и на то, что процесс обновления методик постепенно выведет устаревшие, неэффективные способы контроля из обращения.

#### 4.5.1.2. Период оценки метрологических характеристик

Этот период состоит в анализе контрольного материала (материалов) в течение некоторого числа серий измерений для оценки<sup>82</sup> реальных статистических характеристик методики. Естественно, первый вопрос — какова продолжительность этого периода. Выбор продолжительности периода оценки характеристик анализа, как и при оценке метрологических характеристик методики количественного химического анализа при ее аттестации и формы распределения вообще, представляет собой компромисс между желанием получить возможно более точные оценки и необходимостью быстрее перейти к собственно контролю. Точные рекомендации дать невозможно, однако *в большинстве случаев считается достаточным иметь результаты 20 серий измерений*. Во всяком случае это никак не меньше 15 серий.

Таким образом, период оценки (метрологических) характеристик состоит в анализе в течение 20 серий контрольного материала (материалов) в соответствии с выбранным планом контроля. В результате должны быть оценены метрологические характеристики и проверено соответствие их предъявляемым требованиям. Все это является вариантом *верификации* проверки того, что методика воспроизведена в лаборатории с требуемыми метрологическими характеристиками.

Обозначим результаты, получаемые в период оценки характеристик и после него, как  $X_{ji}$ , где  $j = 1, 2, \dots, M$  — номер серии,  $i = 1, 2, \dots, N$  — номер измерения (табл. 4.2). Как указано выше, чаще всего  $M = 20$ ,  $N = 2$ .

<sup>81</sup> Конечно, это относится в основном к старым методикам. Если в современной методике рекомендованы процедуры «оперативного контроля» по [117], но при этом дается ссылка на возможность контроля стабильности по ГОСТ 5725 [27], проблемы не возникает.

<sup>82</sup> В метрологической литературе используют термин «оценивание».



Таблица 4.2. Результаты, полученные в период оценки характеристик

Номер серии	Номер измерения					
	1	2	...	<i>I</i>	...	<i>N</i>
1	$X_{11}$	$X_{12}$	...	$X_{1i}$	...	$X_{1N}$
2	$X_{21}$	$X_{22}$	...	$X_{2i}$	...	$X_{2N}$
...	...	...	...	...	...	...
<i>J</i>	$X_{j1}$	$X_{j2}$	...	$X_{ji}$	...	$X_{jN}$
...	...	...	...	...	...	...
<i>M</i>	$X_{M1}$	$X_{M2}$	...	$X_{Mi}$	...	$X_{MN}$

Далее начинается собственно контроль стабильности

### Оценка повторяемости (сходимости)

По данным, представленным в табл. 4.2, можно вычислить стандартное отклонение повторяемости  $s_r$  по формуле

$$s_r = \sqrt{\frac{\sum_{j=1}^M \sum_{i=1}^N (X_{ji} - \bar{X}_j)^2}{M(N-1)}}, \quad (4.12)$$

где  $\bar{X}_j = \sum_{i=1}^N \frac{X_{ji}}{N}$ . При  $N = 2$  формула упрощается:

$$s_r = \sqrt{\frac{\sum_{j=1}^M (X_{j1} - X_{j2})^2}{2M}}, \quad (4.13)$$

где  $X_{j1}$  и  $X_{j2}$  — результаты, полученные в  $j$ -й серии.

Полученную величину  $s_r$  сравнивают с нормативом  $\sigma_r$  — приписанным значением стандартного отклонения повторяемости (сходимости), почти всегда приводимым в методике измерений<sup>83</sup>. Если

$$\frac{s_r^2}{\sigma_r^2} < \frac{\chi^2(\bar{P}, f)}{f}, \quad (4.14)$$

где  $\chi^2(\bar{P}, f)$  есть  $\bar{P} = 1 - \alpha$  квантиль  $\chi^2$  — распределения с числом степеней свободы  $f$  (см. Приложение 3), то повторяемость (сходимость) считается удовлетворительной. В отсутствие специальных требований  $\alpha$  принимают равной 0,05. При расчете  $s_r$  по формулам (4.12, 4.13) число степеней свободы равно  $f = M(N - 1)$ .

<sup>83</sup> Часто в методиках приводится величина предела повторяемости (сходимости)  $r$ , в старых методиках обозначаемая как  $d$ . До сих пор иногда это делается с формулировкой «разность двух результатов, получаемых в одной лаборатории, не превышает...». В таких случаях величину  $\sigma_r$  рассчитывают как  $\sigma_r = r/2,77$ .

Оценка повторяемости (сходимости) проводится при оценке статистических характеристик. Однако при контроле по контрольным картам аналогичная оценка может выполняться за любой период, когда велся контроль.

В [27] рекомендован другой способ оценки повторяемости (сходимости), при котором оценка проводится в одной серии измерений. Для такой оценки также необходим контрольный материал: достаточно однородная либо тщательно перемешанная рутинная проба; гомогенизированная смесь нескольких рутинных проб; стандартный образец состава; и так далее. Этот материал анализируют не менее чем  $N = 8 \div 10$  раз в одной аналитической серии и находят оценку стандартного отклонения сходимости  $s_g$  как среднеквадратичное отклонение полученных результатов  $X_1, X_2, \dots, X_i, \dots, X_N$ :

$$s_r = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1}}, \quad (4.15)$$

где  $\bar{X} = \sum_{i=1}^N \frac{X_i}{N}$ . Далее найденную величину  $s_r$  сравнивают с метрологической характеристикой  $\sigma_r$  по формуле (4.14); при этом число степеней свободы  $f = N - 1$ .

Такой способ плохо сочетается с контролем стабильности по контрольным картам, когда контрольный материал исследуется в нескольких сериях и в каждой серии выполняется немного (обычно  $N = 2$ ) измерения. Менее трудоемко, да и более надежно выполнять оценку без «лишних» измерений по формулам (4.12–4.14). Конечно, оценка в одной, специально организованной серии неизбежна, если контроль стабильности не ведется вообще либо при контроле выполняется лишь одно измерение в серии.

В любом случае оценка стандартного отклонения повторяемости (сходимости)  $s_r$  по формуле (4.14) представляет собой проверку гипотезы о равенстве экспериментально полученной дисперсии  $s_r^2$  и заданной нормативным документом, точно известной дисперсии  $\sigma_r^2$ . Такая процедура предполагает «презумпцию невиновности» лаборатории — величина  $s_r$ , несколько превышающая норматив  $\sigma_r$ , признается удовлетворительной<sup>84</sup>.

Если в методике, предусматривающей расчет окончательно приводимого результата как среднего из  $n$  параллельных измерений, дано стандартное отклонение повторяемости среднего значения  $\sigma_{r,\bar{X}}$ , то  $\sigma_r = \sqrt{n}/\sigma_{r,\bar{X}}$ .

#### Оценка промежуточной прецизионности

Имея данные, приведенные в табл. 4.2, стандартное отклонение промежуточной прецизионности  $s_{JP}$  можно рассчитать по формуле

$$s_{JP} = \sqrt{\frac{\sum (\bar{X}_j - \bar{\bar{X}})^2}{M-1} + \left(1 - \frac{1}{N}\right) s_r^2}, \quad (4.16)$$

<sup>84</sup> Вместо формулы (4.14) можно проводить расчет по формуле  $\frac{s_r^2}{\sigma_r^2} < F(\bar{P}, f, \infty)$ , см. также разд. 2.5.1.

$$\text{где } \bar{X}_j = \sum_{i=1}^N \frac{X_{ji}}{N}, \quad \bar{\bar{X}} = \sum_{j=1}^M \frac{\bar{X}_j}{M} \quad ^{85}.$$

Норматив для промежуточной прецизионности в методиках практически всегда отсутствует (см. выше, разд. 3.3.1). Можно, однако, предложить полуколичественный тест, а именно: если  $s_{IP} > 1,5s_r$ , то с высокой вероятностью разброс результатов от серии к серии (который, собственно, и характеризуется промежуточной прецизионностью) чрезмерно велик и желательно принять меры по его уменьшению. И уж конечно, должно выполняться условие  $s_{IP} < \sigma_R$ , где  $\sigma_R$  — стандартное отклонение воспроизводимости, приведенное в методике (норматив)<sup>86</sup>. Если это соотношение не выполняется, безусловно, необходимо принять меры к снижению промежуточной прецизионности.

### Оценка правильности

Как и оценка повторяемости (сходимости), правильность может быть оценена в нескольких сериях (в ходе контроля стабильности) и в одной, специально организованной серии. Если ведется контроль стабильности по контрольным картам, правильность естественно оценивать в нескольких сериях (данные из табл. 4.2). Такая оценка может проводиться только в случае, когда для контроля используется контрольный материал с известным значением: стандартный образец либо аттестованная смесь; образец, предварительно исследованный в авторитетной лаборатории, и т. д. Мы рассмотрим способы с использованием стандартного образца (СО) и образца, предварительно проанализированного другой (авторитетной) лабораторией.

#### Оценка правильности при контроле стабильности с использованием СО.

Проведя измерения в  $M$  сериях (см. табл. 4.2), проверяют гипотезу об отсутствии систематической погрешности. Рассчитывают величину

$$t = \frac{|\bar{\bar{X}} - X_a|}{\sqrt{\frac{s_{\bar{X}_j}^2}{M} + \frac{(\Delta X_a)^2}{3}}}, \quad (4.17)$$

где  $X_a$  и  $\Delta X_a$  — соответственно, аттестованное значение и его погрешность<sup>87</sup>,

$$s_{\bar{X}_j} = \sqrt{\frac{\sum (\bar{X}_j - \bar{\bar{X}})^2}{M-1}}. \text{ Если } t > t(\bar{P}, f), \text{ при выбранном уровне доверительной вероятности } \bar{P} \text{ (обычно 0,975 для двустороннего критерия) и числе степеней свободы}$$

<sup>85</sup> По формуле (ВК6) рассчитывается стандартное отклонение промежуточной прецизионности, включающее  $s_r$ . Поэтому при  $N = 1$  дополнительный вклад повторяемости (сходимости) равен 0, а при  $N = \infty - s_r$ .

<sup>86</sup> Это соотношение предполагает, что величина  $\sigma_R$  относится к единичным измерениям.

<sup>87</sup> Конечно, всегда лучше использовать такой материал, погрешностью аттестации которого согласно принципу пренебрежения малыми погрешностями (см. разд. 2.6.1) можно пренебречь. Однако это не всегда возможно.



$f = M - 1$ , то гипотезу отвергают, что означает наличие систематической составляющей погрешности. В этом случае можно найти оценку систематической погрешности  $\Delta_c = \bar{\bar{X}} - X_a$  и верхнюю границу доверительного интервала для нее

$$\Delta_{c,B} = \left| \bar{\bar{X}} - X_a \right| + \frac{s_{\bar{\bar{X}}_j} t(\bar{P}, f)}{\sqrt{M}} + \Delta X_a. \quad (4.18)$$

Если в методике указана максимально допустимая величина систематической погрешности  $\Delta_{c,max}$ , то необходимо выполнение неравенства

$$\Delta_{c,B} < \Delta_{c,max}. \quad (4.19)$$

В противном случае систематическая погрешность превышает допустимую, и необходимо принять меры по ее уменьшению либо выполнить коррекцию (введение поправки на систематическую погрешность). То же необходимо сделать, если величина  $\Delta_{c,max}$  в методике не приведена, но систематическая погрешность выявлена достоверно.

Оценка правильности при контроле стабильности с помощью контрольного материала, предварительно проанализированного в другой (авторитетной) лаборатории.

Обозначим результаты, полученные при исследовании образца в другой лаборатории, как  $Y_1, Y_2, \dots, Y_i, \dots, Y_K$ . Гипотезу об отсутствии систематической погрешности в этом случае проверяют следующим образом.

Рассчитывают величину

$$t = \frac{|\bar{\bar{X}} - \bar{Y}|}{\sqrt{\frac{s_{\bar{\bar{X}}_j}^2}{M} + \frac{s_Y^2}{K}}}, \quad (4.20)$$

где  $\bar{Y} = \sum_{i=1}^K \frac{Y_i}{K}$ ,  $s_Y^2 = \frac{(Y_i - \bar{Y})^2}{K - 1}$ . Если  $t > t(\bar{P}, f)$ , при уровне доверительной вероятности  $\bar{P}$  (обычно 0,975 для двустороннего критерия) и числе степеней свободы  $f = M + K - 2$ , то гипотезу отвергают, что означает наличие систематической составляющей погрешности. В этом случае оценка систематической погрешности  $\Delta_c = \bar{\bar{X}} - \bar{Y}$ , а верхняя граница доверительного интервала для нее

$$\Delta_{c,B} = \left| \bar{\bar{X}} - \bar{Y} \right| + \Delta X + \Delta Y, \quad (4.21)$$

где  $\Delta X = \frac{s_{\bar{\bar{X}}_j} t(\bar{P}, f)}{\sqrt{M}}$ ,  $\Delta Y = \frac{s_Y t(\bar{P}, f_1)}{\sqrt{K}}$ ,  $f_1 = K - 1$ . Оценка систематической погрешности в случае, когда в методике указана допустимая величина систематической погрешности, проводится по формуле (4.19) и сл.

**Оценка правильности в отдельной серии измерений.** В случае использования для контроля стабильности не аттестованного контрольного материала оценивать правильность приходится в специально организованном эксперименте. При этом можно использовать разные способы: анализ стандартного образца либо образца, проанализированного в авторитетной лаборатории, метод добавок и т. д. Рассмотрим способы с использованием СО и образца, проанализированного другой (авторитетной) лабораторией.

**Оценка правильности в одной серии с использованием СО.** Стандартный образец анализируют несколько ( $N$ ) раз в одной серии. Далее проверяют гипотезу об отсутствии систематической погрешности следующим образом. Рассчитывают величину

$$t = \frac{|\bar{X} - X_a|}{\sqrt{\frac{s_r^2}{N} + \frac{(\Delta X_a)^2}{3}}}, \quad (4.22)$$

где  $X_a$  и  $\Delta X_a$  — соответственно, аттестованное значение СО и его погрешность. Если  $t > t(\bar{P}, f)$ , при выбранном уровне доверительной вероятности  $\bar{P}$  (обычно 0,975 для двустороннего критерия) и числе степеней свободы  $f = N - 1$ , то гипотезу отвергают, что означает наличие систематической составляющей погрешности. В этом случае оценивают полную систематическую погрешности  $\Delta_c = \bar{X} - X_a$  и верхнюю границу доверительного интервала для нее

$$\Delta_{c,v} = |\bar{X} - X_a| + \frac{s_r t(\bar{P}, f)}{\sqrt{N}} + \Delta X_a. \quad (4.23)$$

Если в методике указана максимально допустимая величина систематической погрешности  $\Delta_{c,max}$ , то необходимо выполнение неравенства (4.19) и сл.

**Оценка правильности в одной серии с помощью образца, проанализированного в другой (авторитетной) лаборатории.**

Такой образец анализируют в одной серии несколько ( $N$ ) раз. Как и раньше, обозначим результаты, полученные при исследовании образца в другой лаборатории, как  $Y_1, Y_2, \dots, Y_r, \dots, Y_K$ . Проверяют гипотезу об отсутствии систематической погрешности, рассчитывая величину

$$t = \frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{s_r^2}{N} + \frac{s_Y^2}{K}}}. \quad (4.24)$$

Если  $t > t(\bar{P}, f)$  при выбранном уровне доверительной вероятности  $\bar{P}$  (обычно 0,975 для двустороннего критерия) и числе степеней свободы  $f = N + K - 2$ , то гипотезу отвергают, что означает наличие систематической составляющей погрешности. В этом случае оценка систематической погрешности  $\Delta_c = \bar{X} - \bar{Y}$ , а верхняя граница доверительного интервала для нее

$$\Delta_{c,b} = |\bar{X} - \bar{Y}| + \Delta X + \Delta Y, \quad (4.25)$$

где  $\bar{Y} = \frac{\sum Y_i}{K}$ ,  $\Delta X = \frac{s_{r,t}(\bar{P}, f)}{\sqrt{N}}$ ,  $\Delta Y = \frac{s_{r,Y}t(\bar{P}, f_1)}{\sqrt{K}}$ ,  $f_1 = K - 1$ . Оценка систематической погрешности в случае, когда в методике указана допустимая величина систематической погрешности, проводится по неравенству (4.19) и сл.

Если контроль стабильности планируется вести по неаттестованному материалу, проверку правильности на этапе оценки метрологических характеристик можно проводить любым из вышеописанных способов. Если при этом используется стандартный образец, можно делать это:

- в отдельной серии измерений (см. формулу 4.23 и сл.);
- анализируя СО вместе с контрольным материалом в течение всего периода оценки характеристик, желательно по два измерения в серии, с оценкой правильности по формулам (4.17 и сл.).

Последнее надежнее.

#### **Фильтрация результатов измерений в ходе оценки характеристик**

Часто (хотя и не всегда) перед собственно оценкой характеристик выявляют и отбрасывают грубые промахи. При этом используются два основных статистических теста:  $\zeta$ -критерий [23] (он же «критерий Граббса» в [27]) и критерий Кохрена.

$\zeta$ -критерий (критерий Граббса) похож на правило грубой ошибки (см. разд. 2.4.2) и предназначен для выявления грубого промаха в совокупности результатов измерений, имеющих простейшую структуру  $X_1, X_2, \dots, X_j, \dots, X_M$ . В применении к нашим данным (табл. 4.2)  $X_j = \sum_{i=0}^N X_{ji} / N$ . Пусть величины  $X_j$  расположены по возрастанию. Тогда  $\zeta$ -критерий (критерий Граббса) рассчитывается как  $\zeta = (X_M - \bar{X}) / s$ , если подозрительно выделяется  $X_M$ , либо как  $\zeta = (\bar{X} - X_1) / s$ , если подозрительно выделяется  $X_1$ . Если  $\zeta > \zeta(\alpha, M)$ , где  $\zeta(\alpha, M)$  – критическое значение  $\zeta$ -статистики при вероятности ошибки 1-го рода  $\alpha$ , то подозрительно выделяющееся значение считают грубым промахом и удаляют из дальнейших расчетов. В [27] соответствующая статистика обозначена как  $G$ . Обычно выбирают  $\alpha = 0,05$  или  $\alpha = 0,01$  (таблицы критических значений имеются в [22, 27]).

Использование критерия Кохрена (Кочрена; Cochren) для отброса сильно выделяющихся серий выборок) описано нами в разд. 2.5.2.

Оба критерия предполагают нормальное распределение результатов как внутри серий, так и между ними. В общем-то, можно проводить фильтрацию последовательно несколько раз, однако надо очень хорошо подумать, прежде чем отбрасывать более чем одну серию из 20 – это может привести к чрезмерному цензурированию выборки и нереальным оценкам метрологических характеристик.

Во многих случаях, если какую-либо серию отбрасывают (отфильтровывают), вместо нее выполняют еще одну.

Если по завершении стадии оценки метрологических характеристик была проведена фильтрация данных по  $\zeta$ -критерию (критерию Граббса) и критерию Кохрена и затем корректно велся контроль стабильности по контрольным картам средних



и расхождений (размахов) (см. ниже), в дальнейшем применение этих критериев обычно не выявляет достоверных промахов.

Проверка соответствия метрологических характеристик, достигнутых лабораторией, нормативным проводится по окончании периода оценки характеристик. Однако понятно, что в реальности опасно дожидаться окончания этого достаточно длительного периода — можно оказаться у разбитого корыта, если какая-то характеристика окажется неудовлетворительной. Поэтому естественно проверять качество анализа несколько раз по мере набора данных — например, после набора 4–5; 8–10 и т. д. серий — и продолжать работать только при удовлетворительных результатах. В противном случае набор данных останавливают и принимают меры по повышению качества измерений.

Далее переходят к собственно контролю стабильности.

#### *4.5.1.3. Контроль стабильности с помощью контрольных карт при наличии контрольного материала*

Когда период оценки характеристик успешно завершен, рассчитывают соответствующие контрольные пределы, строят контрольные карты и начинают контроль стабильности. При этом для построения контрольных карт используют данные, полученные за этот период.

При контроле стабильности можно отдельно и разными способами контролировать а) точность анализа; б) повторяемость (сходимость) и в) промежуточную прецизионность результатов анализа. Естественно, сходимость может контролироваться только в случае проведения нескольких измерений для данного контрольного материала в каждой серии. Раздельный контроль более эффективен и позволяет легче выявлять источники погрешностей.

Основным инструментом контроля стабильности являются контрольные карты, которые впервые были предложены Шухартом для контроля качества производства промышленных изделий [119, 120]. В аналитической химии эти карты были впервые применены в работе [121], поэтому здесь их иногда называют «карты Леви — Дженнингса». Контрольные карты в химическом анализе представляют собой графики, на которых по горизонтальной оси откладывают порядковый номер серии измерений (либо дата исследования), а по вертикальной — концентрацию определяемого вещества в контрольном материале, разницу между результатами измерений, кумулятивные суммы и т. д. На эти карты наносятся в виде горизонтальных прямых различные «пределы», и выход графика за эти пределы свидетельствует о нестабильности процесса измерений. Можно сказать, что по оси ординат на контрольных картах откладываются величины, которые должны были бы быть постоянными в отсутствие погрешностей измерений.

Условием использования контрольных карт является то, что при переходе от одной серии к другой должны меняться одни и те же факторы. В химическом анализе чаще всего изменяются факторы «время» и «калибровка», а «лаборант» и «оборудование» остаются постоянными, однако в крупных лабораториях могут меняться любые факторы (особенно при сменной работе).

**Использование контрольных карт.** Контрольные карты используются для выявления случаев выхода процесса измерений из статистически контролируемого состояния (см. выше). Если статистические характеристики процесса стабильны, разброс точек на контрольных картах является чисто случайным. Отклонения от случайного характера расположения точек на картах, свидетельствующие о выходе процесса из-под контроля, могут носить различный характер: выбросы, скачки, дрейфы и т. д.

Важнейший способ использования контрольных карт — их визуальное рассмотрение. Глазом часто удается выявить такие закономерные изменения аналитического процесса, собственно, и свидетельствующие о выходе измерения из-под контроля либо об опасных тенденциях, которые с трудом обнаруживаются при статистическом анализе. Однако при рутинной работе больше опираются на формальные признаки, указывающие на выход процесса измерений из-под контроля.

Для диагностики выхода методики количественного химического анализа из статистически контролируемого состояния используют различные правила. Можно выделить два вида таких правил.

*Тревожные (или «предупредительные»)<sup>88</sup> признаки* свидетельствуют лишь о достаточно большой вероятности снижения качества анализа. Возникновение такого события указывает на большую вероятность того, что в ближайшем будущем процесс анализа выйдет из-под контроля. Появление тревожного признака (признаков) не требует остановки измерений и повторного исследования уже проанализированных проб, но должно привлечь внимание исполнителей и стать сигналом для принятия профилактических мер.

*Контрольные признаки* (они же *контрольные критерии, границы регулирования* и др.) свидетельствуют о достоверном<sup>89</sup> выходе процесса измерений из подконтрольного состояния. При их появлении лучше всего приостановить анализ, выявить и устранить причины сбоя и повторно исследовать пробы из последней, неудачной аналитической серии. Сведения о тревожных и контрольных признаках для разных контрольных карт приведены, в частности, в [120].

Основные виды контрольных карт, используемые для контроля стабильности для случая, когда лаборатория имеет подходящий для данной методики контрольный материал, можно проиллюстрировать с помощью данных, приведенных в табл. 4.3.

В табл. 4.3 даны результаты определения водорода в стандартном образце ацетанилида  $C_8H_9NO$  (контрольный материал), полученные при контроле качества исследований элементного состава вновь синтезируемых веществ. Концентрацию водорода наряду с содержанием других элементов определяли методом газовой хроматографии после сжигания пробы в динамической вспышке на C, H, N, S — анализаторе EA1108 фирмы «Carlo Erba Instr.». Водород был выбран для контроля стабильности как наиболее «капризный» из четырех определяемых совместно элементов (см. выше, разд. 4.3.1.1).

<sup>88</sup> Тревожные признаки свидетельствуют о выходе методики из статистически контролируемого состояния с доверительной вероятностью  $\bar{P} \approx 0,95$ .

<sup>89</sup> Рассмотренные ниже контрольные признаки выявляют выход методики из-под контроля с  $\bar{P} > 0,99$ .

Таблица 4.3. Результаты измерения концентрации водорода в стандартном образце ацетанилида за октябрь–ноябрь 2002 г. (в процентах по массе)

Дата анализа	Данные наблюдений		Среднее значение $\bar{X}_j = \frac{X_1 + X_2}{2}$	$\bar{\delta}_j = \bar{X}_j - \mu$	Расхождение (размах) $ X_1 - X_2 $	Текущее расхождение $w_j =  \bar{X}_{j+1} - \bar{X}_j $	CUSUM	Примечание
	$X_1$	$X_2$						
03.10.02	7,20	7,08	7,140	+0,430	0,12	0,325		
07.10.02	6,81	6,82	6,815	+0,105	0,01	0,285		
10.10.02	6,54	6,52	6,530	-0,180	0,02	0,375		
11.10.02	6,18	6,13	6,155	-0,555	0,05	0,790		
14.10.02	6,92	6,97	6,945	+0,235	0,05	0,200		
15.10.02	6,79	6,70	6,745	+0,035	0,09	0,185		
17.10.02	6,63	6,49	6,560	-0,150	0,14	0,100		
18.10.02	6,75	6,57	6,660	-0,050	0,18	0,080		
21.10.02	6,65	6,51	6,580	-0,130	0,14	0,095		
22.10.02	6,68	6,67	6,675	-0,035	0,01	0,430		
24.10.02	7,03	7,18	7,105	+0,395	0,15	0,715		
25.10.02	6,28	6,50	6,390	-0,320	0,22	0,165		
28.10.02	6,17	6,28	6,225	-0,485	0,11	0,665		
29.10.02	6,91	6,87	6,890	+0,180	0,04	0,410		
31.10.02	6,52	6,44	6,480	-0,230	0,08	0,305		
01.11.02	6,76	6,81	6,785	+0,075	0,05	0,005		
04.11.02	6,73	6,86	6,790	+0,080	0,13	0,155		
05.11.02	6,67	6,60	6,635	-0,075	0,07	0,005		
08.11.02	6,71	6,55	6,630	-0,080	0,16	0,115		
11.11.02	6,57	6,46	6,515	-0,195	0,11	0,155		Уточнены параметры КК*
12.11.02	6,22	6,50	6,360	-0,350	0,28	0,015	-0,175	7** 4D на КК средних значений; R(2σ) на КК расхождений
14.11.02	6,41	6,28	6,345	-0,365	0,13	0,095	-0,366	7 4D, 2(1σ) на КК средних значений; R(2σ) на КК расхождений
15.11.02	6,44	6,06	6,250	-0,460	0,38	0,160	—	///R(3σ) на КК расхождений
18.11.02	6,36	6,46	6,410	-0,300	0,10	0,200	-0,491	
19.11.02	6,40	6,02	6,210	-0,500	0,38	0,015	—	///R(3σ) на КК расхождений



Таблица 4.3 (окончание)

Дата анализа	Данные наблюдений		Среднее значение $\bar{X}_j = \frac{X_1 + X_2}{2}$	$\bar{\delta}_j = \bar{X}_j - \mu$	Расхождение (размах) $ X_1 - X_2 $	Текущее расхождение $w_j =  \bar{X}_{j+1} - \bar{X}_j $	CUSUM	Примечание
	$X_1$	$X_2$						
21.11.02	6,10	6,29	6,195	-0,515	0,19	0,140	-0,831	$7(\bar{X})$ на КК средних значений
22.11.02	6,00	6,11	6,055	-0,555	0,11	0,330	-1,312	///5.1s на карте CUSUM, $7(2\sigma)$ , $2(1\sigma)$ , $7(\bar{X})$ на КК средних значений
23.11.02	5,58	5,85	5,725	-0,985	0,27	1,350	—	///1(3 $\sigma$ ), $2(2\sigma)$ на КК средних значений;
26.11.02	6,99	7,16	7,075	+0,365	0,17	0,370	+0,285	///R(3 $\sigma$ ) на КК текущих расхождений
28.11.02	6,70	6,71	6,705	-0,005	0,01	0,240	+0,200	
29.11.02	6,99	6,90	6,945	+0,235	0,09		+0,355	

\* КК — контрольная карта (карты). \*\* ? — «тревожный признак». \*\*\* !!! — «контрольный признак»

**Пояснения к таблице**  
 Аттестованное значение для содержания водорода  $\mu$  в стандартном образце ацетилинда 6,71 %.  
 До 11.11.02 велся набор данных для оценки параметров контрольных карт, после чего начался собственно контроль. 11.11.02 были оценены следующие параметры контрольных карт:  
 Стандартное отклонение  $s_x = 0,255\%$ , общее среднее  $\bar{X} = 0,663\%$ .  
 Среднее расхождение (размах)  $R = 0,0965\%$ . Среднее текущее расхождение  $\bar{w} = 0,285\%$ .  
 До 11.11.02 контроль по данному материалу не велся (период оценки характеристик). В это время контроль стабильности велся с помощью другого стандартного образца, и его результаты были удовлетворительными.  
 12.11 были получены результаты для 20 серий и рассчитаны параметры контрольных карт. После этого начался собственно контроль с использованием ацетилинда.  
 12.11 были выявлены «тревожные признаки» на контрольных картах средних значений и расхождений (размахов). Они были оставлены без внимания.  
 14.11 вновь появились тревожные признаки на контрольных картах средних значений и расхождений (размахов). Проведено обычное техническое обслуживание автоанализатора.  
 15.11 появился «контрольный признак» R(3 $\sigma$ ) на карте расхождений (размахов). Проверена работа системы контроля температуры сгорания пробы, и 18.11 routine пробы были проанализированы повторно.  
 19.11 вновь выявлен «контрольный признак» R(3 $\sigma$ ) на карте расхождений. Проверена работа детектора (катарометр).  
 21.11 выявлен «тревожный признак»  $7(\bar{X})$  на контрольной карте средних значений (указание на возможность систематической погрешности). Пробы от 19.11 проанализированы повторно.  
 22.11 появился «контрольный признак 5.1s» (выход за предел 5.1s) на контрольной карте CUSUM (наряду с тревожными признаками на других контрольных картах). Анализ был остановлен, вновь проведены стандартные профилактические мероприятия. Тем не менее 23.11 выявлены контрольные признаки на картах средних значений и текущих расхождений. После этого исследованием были остановлены, прибор разобран. Проведена замена треснувшего кварцевого реактора.  
 26.11 исследование были возобновлены и, судя по результатам контроля качества, все пришло в норму (контрольный признак R(3 $\sigma$ ) появился 26.11 на карте текущих расхождений из-за скачка при переходе процесса измерений в нормальное состояние). Пробы от 22.11 проанализированы повторно.  
 В дальнейшем процесс измерений находился в статистически контролируемом состоянии.  
 Заметим, что, как и следовало ожидать, на карте CUSUM проблема (возникающая систематическая погрешность) была выявлена раньше, чем на карте средних значений. Таким образом, с помощью контрольных карт были выявлены аппаратные проблемы и предотвращена выдача заказчикам неправильных результатов исследований (хотя, возможно, решительные меры были предприняты с опозданием).

В те дни, когда выполнялись исследования, лаборатория анализировала от 10 до 20 рутинных проб органических соединений в условиях повторяемости (сходимости), то есть одна серия в день. Среди этих проб случайным образом располагали две пробы стандартного образца ацетанилида фирмы «Carlo Erba Instr.», по результатам анализа которых и велся контроль стабильности измерений. Результаты исследований стандартного образца за период октябрь–ноябрь 2002 г., с начала использования данного контрольного материала, и приведены в табл. 4.3. При этом в каждой строке расположены результаты для  $j$ -й серии измерений, а по столбцам – следующие данные:

- дата выполнения измерений;
- результаты двух измерений концентрации водорода в стандартном образце  $X_1$  и  $X_2$ , выраженные в относительных процентах;
- среднее значение для результатов  $j$ -й серии измерений  $\bar{X}_j = (X_1 + X_2) / 2$ ;
- оценка систематической погрешности  $\hat{\delta}_j = \bar{X}_j - \mu$  ( $\mu$  – принятое опорное значение концентрации в контрольном материале). В нашем случае это содержание водорода в стандартном образце ацетанилида, равное 6,71%;
- размах (расхождение), то есть разность между максимальным и минимальным значениями для данной серии  $R = |X_1 - X_2|$ . В [27] эта величина называется «расхождение», хотя обычно в отечественной литературе используется термин «размах», см. разд. 2.2.2;
- текущее расхождение  $w_i = |\bar{X}_{j+1} - \bar{X}_j|$ ;
- величина кумулятивной суммы  $CUSUM$  (способ ее расчета см. ниже);
- примечания, в которых указаны виды нестабильности в соответствии с правилами Вестгарда (см. ниже).

Наиболее эффективны и чаще всего используются следующие виды контрольных карт.

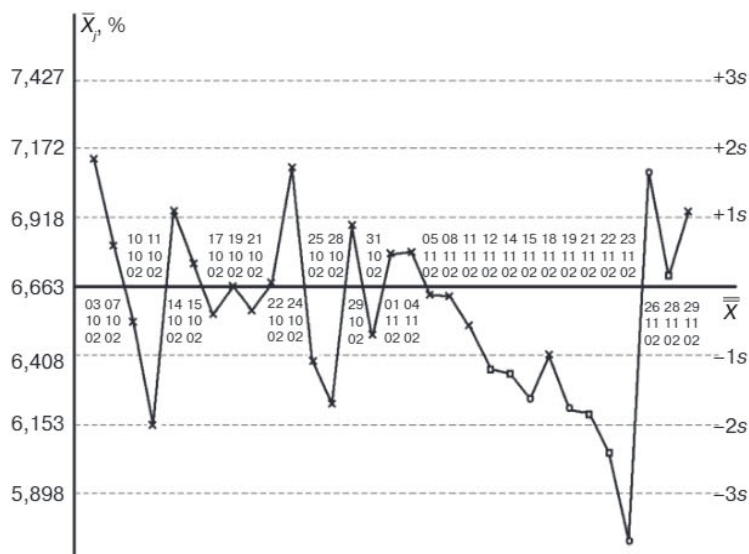


Рис. 4.3. Контрольная карта средних значений. Данные из табл. 4.3

**Контрольные карты средних значений**, показанные для данных из табл. 4.3 на рис. 4.3<sup>90</sup>, наиболее универсальны, весьма эффективны и используются для выявления любых видов статистически значимых проявлений нестабильности (разовые выбросы, небольшие, но стабильные систематические погрешности, дрейфы и т. д.). На этих картах по оси абсцисс откладывают номер серии (дату анализа), а по оси ординат — средние значения, полученные для контрольного материала в  $j$ -й серии  $\bar{X}_j$ <sup>91</sup>. Конечно, если в каждой аналитической серии выполняется только одно контрольное измерение, его результат и есть среднее. В виде горизонтальных линий наносят общее среднее значение  $\bar{\bar{X}} = \sum \frac{\bar{X}_j}{M}$  и «контрольные пределы»  $(\bar{\bar{X}} \pm 1s)$ ,  $(\bar{\bar{X}} \pm 2s)$ ,  $(\bar{\bar{X}} \pm 3s)$ , где  $s \equiv s_{\bar{X}_j} = \sqrt{\sum (\bar{X}_j - \bar{\bar{X}})^2 / (M - 1)}$  — стандартное отклонение величин  $\bar{X}_j$ . Величины  $\bar{\bar{X}}$  и  $s$  оцениваются по результатам исследования контрольного материала, полученным в период оценки характеристик (см. выше) и, обычно, уточняются каждый раз после набора очередных 20–30 серий. Как указано выше, *в каждой серии чаще всего проводится два измерения для каждого контрольного материала (как и в нашем примере), то есть при расчете  $\bar{X}_j$  для каждой  $j$ -й серии усредняются два результата измерений*.

В оригинальной работе [114] был предложен один тревожный признак  $1(2s)$  и один контрольный признак  $1(3s)$ , то есть выход результата последнего контрольного измерения за границы  $(\bar{\bar{X}} \pm 2s)$  и  $(\bar{\bar{X}} \pm 3s)$  соответственно. Однако быстро выяснилось, что этих правил недостаточно. Ниже изложен набор правил, рекомендуемый в настоящее время при работе с контрольными картами средних значений [122–124].

Следующие «*контрольные признаки*» рассматриваются в случае появления тревожного признака  $1(2s)$ :

- $1(3s)$  — последняя точка лежит выше контрольного предела  $(\bar{\bar{X}} + 3s)$  или ниже контрольного предела  $(\bar{\bar{X}} - 3s)$ . Этот признак используется в [27];
- $2(2s)$  — две последовательных точки, включая последнюю, лежат выше контрольного предела  $(\bar{\bar{X}} + 2s)$  или ниже контрольного предела  $(\bar{\bar{X}} - 2s)$ ;
- $D(4s)$  — разность между последней и предыдущей точками по абсолютной величине превышает  $4s$ ;
- $4(1s)$  — четыре последовательных точки, включая последнюю, лежат выше контрольного предела  $(\bar{\bar{X}} + 1s)$  или ниже контрольного предела  $(\bar{\bar{X}} - 1s)$ ;
- $10(\bar{\bar{X}})$  — десять последовательных точек, включая последнюю, лежат выше (или ниже) среднего значения  $\bar{\bar{X}}$ .

Если хотя бы один из этих признаков выявлен, то процесс измерений вышел из-под контроля. Необходимо повторить измерения, выполненные в последней серии, после выявления и устранения источника погрешности. В результате получается алгоритм, показанный на рис. 4.4 («правила Вестгарда»)<sup>92</sup>.

<sup>90</sup> Контрольные карты созданы с помощью программы QControl, версия 5.5.

<sup>91</sup> Здесь и ниже используются обозначения из табл. 4.3.

<sup>92</sup> В случае контроля с использованием двух контрольных материалов наряду с описанными можно использовать и правила, учитывающие события на двух картах средних значений: появление признаков  $1(2s)$  и  $2(1s)$  на обеих картах в одной серии.



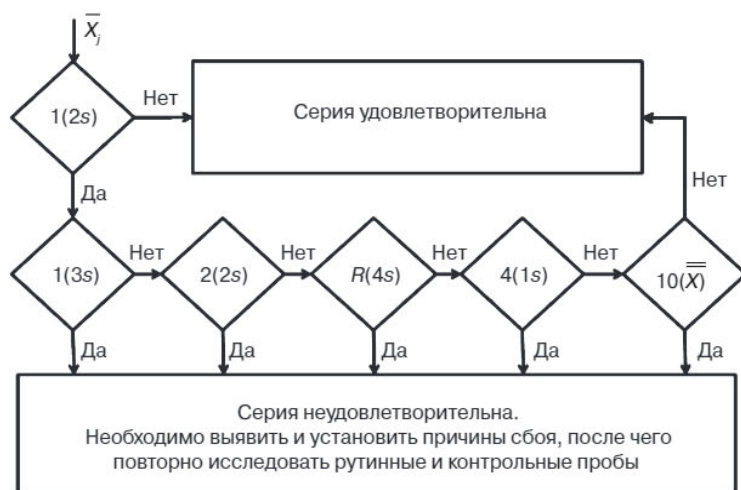


Рис. 4.4. Алгоритм Вестгарда

В качестве «тревожных признаков» используются следующие правила:

- 1(2s) – последняя (то есть вновь введенная) точка лежит выше контрольного предела ( $\bar{\bar{X}} + 2s$ ) или ниже контрольного предела ( $\bar{\bar{X}} - 2s$ ). Этот признак используется и в [27];
- 2(1s) – две или три последовательных точки, включая последнюю, лежат выше контрольного предела ( $\bar{\bar{X}} + 1s$ ) или ниже контрольного предела ( $\bar{\bar{X}} - 1s$ );
- 7( $\bar{\bar{X}}$ ) – семь или более последовательных точек, включая последнюю, лежат выше (или ниже) среднего значения  $\bar{\bar{X}}$ ;
- 4D – каждая из четырех и более последовательных точек, включая последнюю, лежит выше (положительный дрейф) или ниже (отрицательный дрейф) предыдущей.

Появление любого из этих признаков указывает на достаточно высокую (~0,95) вероятность выхода процесса измерений из-под контроля.

Этот алгоритм имеет ряд преимуществ:

- он эффективно выявляет даже не слишком большие систематические погрешности;
- обеспечивает небольшую вероятность ошибочного заключения о выходе измерения из-под контроля;
- позволяет не просто установить факт выхода измерения из-под контроля, но и дает «наводку», где искать проблему. Если сработали правила 1(3s) или D(4s) – скорее всего произошел скачек, какое-то событие типа «лопнула пружина дозатора» или «ошиблись при приготовлении маточного градуировочного раствора». Если же выявлены события 4(1s) или 10( $\bar{\bar{X}}$ ), то это, вероятно, говорит о дрейфе либо о небольшой по величине, но постоянной систематической погрешности (например, происходит испарение градуировочных растворов). Конечно, это лишь подсказка; в любом случае определение и устранение причин выхода методики из-под контроля – творческая задача, требующая квалификации и опыта.

*Альтернативная стратегия использования контрольных карт средних значений, основанная на учете требований к качеству продукции*

Использование описанного выше алгоритма (правил Вестгарда) оптимально в случаях, когда лаборатория обязана обеспечить метрологические характеристики методики не хуже приписанных этой методике характеристик, и реальные характеристики не слишком отличаются от требуемых. Такой подход положен в основу нормативных документов [6, 27, 117].

Недавно был развит другой подход [125,126], при котором используемые признаки (наборы признаков), свидетельствующие о выходе процесса измерений из-под контроля, выбираются с учетом как результатов, полученных при оценке метрологических характеристик (стандартное отклонение, смещение), так и допустимой общей погрешности измерений (выбираемой в данном случае из соображений производственной необходимости). Делается это следующим образом.

Оценивается допустимая («критическая») величина систематической погрешности измерений  $\Delta SE_{\text{крит}}$ , которая может привести к некорректной оценке качества продукции с риском 5% (данный подход был предложен для клинико-диагностических исследований и, соответственно,  $\Delta SE_{\text{крит}}$  характеризует риск неправильной постановки диагноза). Используется соотношение

$$\Delta SE_{\text{крит}} = [(TE_a - bias_{\text{изм}})/s_{\text{изм}}] - 1,65, \quad (4.26)$$

где  $TE_a$  — допустимая общая погрешность измерений<sup>93</sup>,  $bias_{\text{изм}}$  — экспериментально полученное смещение метода (оценка систематической погрешности),  $s_{\text{изм}}$  — стандартное отклонение результатов измерений и 1,65 — величина  $z$  — индекса, отсекающая 5% площади на «хвосте» функции плотности нормального распределения. С учетом реальных характеристик измерения и требований к нему (величин  $\Delta SE_{\text{крит}}$  и  $TE_a$  и результатов ранее проведенного методом Монте-Карло моделирования) выбираются контрольный признак (либо их сочетание), а также число контрольных измерений, обеспечивающие достаточно надежное выявление неудовлетворительных результатов при минимальных затратах. При этом наборы правил оказываются более «мягкими» по сравнению с алгоритмом Вестгарда (рис. 4.4), что позволяет реже вмешиваться в процесс анализа.

Такой подход имеет следующие недостатки:

- оценка систематической погрешности в реальности зачастую затруднительна (в химическом анализе из-за матричных эффектов, приводящих к разным смещениям при использовании разных стандартных образцов);
- появляется необходимость изменять число контрольных измерений при изменении метрологических характеристик методики, что может потребовать нового набора статистики для контрольных карт.

Из-за этого предложенный в [125,126] подход к использованию карт средних значений пока не получил широкого применения.

<sup>93</sup> Величина  $TE_a$  определяется требованиями заказчика исследований, которые, в свою очередь, связаны с тем, как используются результаты измерений.

**Контрольные карты размахов (расхождений)** предназначены для контроля стабильности стандартного отклонения повторяемости (сходимости)  $s_r$ . Такой контроль возможен только в случае, когда в каждой серии (в условиях повторяемости согласно [27]) выполняется несколько измерений для контрольного материала. На контрольную карту размахов (расхождений) по горизонтальной оси наносятся номера серии  $j$ , а по вертикальной — величины размахов ( $X_{j,\max} - X_{j,\min}$ ) — разности между максимальным и минимальным результатами измерений контрольного материала, выполненными в  $j$ -й серии. На рис. 4.5 показана контрольная карта размахов (расхождений) для данных из табл. 4.3. Поскольку в нашем примере рассмотрен случай двух измерений в каждой серии ( $N = 2$ ), по оси ординат отложены величины  $|X_1 - X_2|$ . Как указывалось выше, число измерений в серии может быть и больше (что повышает надежность контроля), однако и стоимость контроля при этом растет<sup>94</sup>. На карту расхождений (размахов) в виде горизонтальных линий наносят среднюю линию и верхние контрольные пределы « $+2s$ » и « $+3s$ » («предел предупреждения» и «предел действия»). Эти величины рассчитывают, умножая оценку стандартного отклонения  $s_r$  (см. выше) на коэффициенты, приведенные в табл. 4.4<sup>95</sup>.

**Таблица 4.4.** Коэффициенты для расчета положения средней линии, контрольных пределов « $+2s$ » и « $+3s$ » для контрольных карт размахов (расхождений) [23, 27, 120]

Число измерений в серии $N$	Коэффициент для расчета положения средней линии ( $d_2$ )	Коэффициент для расчета (верхнего) контрольного предела « $+2s$ »	Коэффициент для расчета (верхнего) контрольного предела « $+3s$ »
2	1,128	2,834	3,686
3	1,693	3,469	4,358
4	2,059	3,819	4,698
5	2,326	4,054	4,918

Заметим, что обозначения « $+2s$ » и « $+3s$ » не имеют прямого отношения к стандартному отклонению и говорят только о том, что эти линии играют ту же роль, что и контрольные пределы ( $\bar{X} \pm 2s$ ), ( $\bar{X} \pm 3s$ ) на контрольных картах средних значений.

Как «тревожный признак» обычно рассматривают правило:

- $R$ (« $2s$ ») — расхождение (размах) в последней (оцениваемой) серии больше верхнего предела предупреждения « $+2s$ ».

В качестве «контрольных признаков» для карт размахов (расхождений) используют следующие правила:

<sup>94</sup> Существуют и контрольные карты стандартных отклонений, используемые для тех же целей, что и контрольные карты размахов. На них вместо размахов для каждой серии наносятся стандартные отклонения результатов измерений, рассчитанные для данной серии. Однако при числе параллельных измерений менее 8 (а в лабораториях больше измерений в серии не бывает) эффективнее использовать расхождение (размах)  $w$ .

<sup>95</sup> Контрольные карты размахов имеют как верхние, так и нижние контрольные пределы (правда, нижний контрольный предел « $-2s$ », соответствующий доверительной вероятности  $P \approx 0,99$ , отличен от нуля лишь при числе параллельных измерений больше трех), однако на практике этот факт в большинстве случаев игнорируют и используют только верхние контрольные пределы.



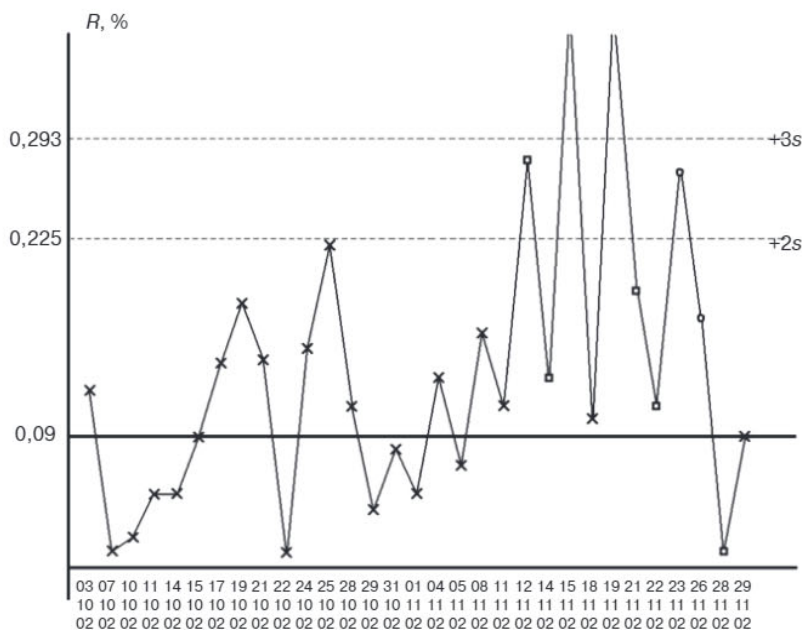


Рис. 4.5. Контрольная карта размахов (расхождений). Данные из табл. 4.3

- $R(«3s»)$  — расхождение (размах) в последней серии больше верхнего предела действия « $+3s$ »;
- $2R(«2s»)$  — расхождение (размах) для двух последних серий больше верхнего предела предупреждения « $+2s$ ».

**Контрольные карты кумулятивных сумм (CUSUM)** предназначены для быстрого выявления небольших по величине постоянных систематических погрешностей и дрейфов. По вертикали на них откладывают кумулятивные суммы  $C_1, C_2, \dots, C_k, \dots$ , которые рассчитывают по формулам:

$$C_1 = (\bar{X}_1 - \bar{\bar{X}}), C_2 = C_1 + (\bar{X}_2 - \bar{\bar{X}}), \dots, C_k = C_{k-1} + (\bar{X}_k - \bar{\bar{X}}), \dots \quad (4.27)$$

Здесь  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k, \dots$  — средние значения результатов анализа контрольного материала, полученные начиная с момента начала расчета кумулятивных сумм,  $\bar{\bar{X}}$  — общее среднее, полученное к этому моменту в предшествующем периоде. Контрольные карты CUSUM, как и другие карты, начинают вести по окончании периода оценки характеристик. В этом случае  $\bar{X}_1$  — среднее значение, полученное в первой серии после завершения этого периода,  $\bar{\bar{X}}$  — общее среднее, полученное за период оценки характеристик (см. выше). На контрольной карте CUSUM ось  $X$  соответствует нулевому значению. Таким образом, каждая точка карты CUSUM представляет собой сумму отклонения текущего результата измерения (среднего значения в серии) от среднего и всех предыдущих отклонений. Отклонение от общего среднего накапливается, что и дало название этому виду контрольных карт.

Если на карту CUSUM наносятся собственно величины  $C_k$  (как это делалось, когда их изобрели), то при их визуальном рассмотрении основное внимание уде-

ляют наклону линии, соединяющей точки карты. Если возникает даже небольшая, но постоянная систематическая погрешность, то эта линия будет «ползти» либо вверх (при положительной погрешности), либо вниз — при отрицательной. Для визуального изучения оптимален масштаб, при котором при расстоянии между точками по горизонтали одна единица такое же расстояние по вертикали равно  $2s$  (здесь, как для карт средних значений,  $s \equiv s_{\bar{x}_j}$ ). Количественное отслеживание угла

наклона часто вели с помощью специальных угловых шаблонов, что не слишком удобно. Карты *CUSUM* детально обсуждаются, например, в [127].

В настоящее время используются более удобные представления карт *CUSUM*. Приведем наиболее эффективный, на наш взгляд, способ, близкий к приведенному в [27].

До тех пор, пока очередной результат  $\bar{X}_k$  не выходит за пределы  $(\bar{\bar{X}} \pm 0,5s)$ , кумулятивные суммы не рассчитываются и не наносятся на контрольную карту. В момент выхода за пределы  $(\bar{\bar{X}} \pm 0,5s)$  начинают рассчитывать кумулятивные суммы. При этом расчет проводят исходя из того, верхнюю или нижнюю границу пересекла кривая, соединяющая точки на карте средних значений. Если это  $(\bar{\bar{X}} + 0,5s)$ , то при расчете кумулятивных сумм из каждого следующего значения вычитают не  $\bar{X}$ , а  $(\bar{X} + 0,5s)$ ; если пересечена граница  $(\bar{\bar{X}} - 0,5s)$ , то вычитают  $(\bar{X} - 0,5s)$ . Если после этого очередная кумулятивная сумма окажется больше  $+5,1s$  в первом случае или ниже  $-5,1s$  во втором, то систематическая погрешность считается достоверно выявленной («контрольный признак  $5,1s$ »). Если же вместо этого кривая, соединяющая кумулятивные суммы, пересечет нулевую линию, их расчет прекращают впредь до нового выхода очередного результата за пределы  $(\bar{\bar{X}} \pm 0,5s)$ .

Кумулятивные суммы, приведенные в табл. 4.3, рассчитаны по этому алгоритму. Соответствующая контрольная карта — на рис. 4.6.

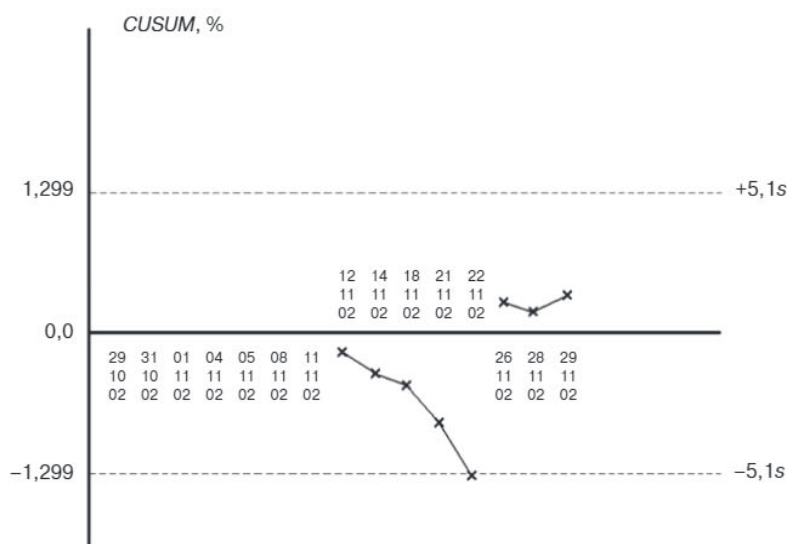


Рис. 4.6. Контрольная карта кумулятивных сумм. Данные из табл. 4.3

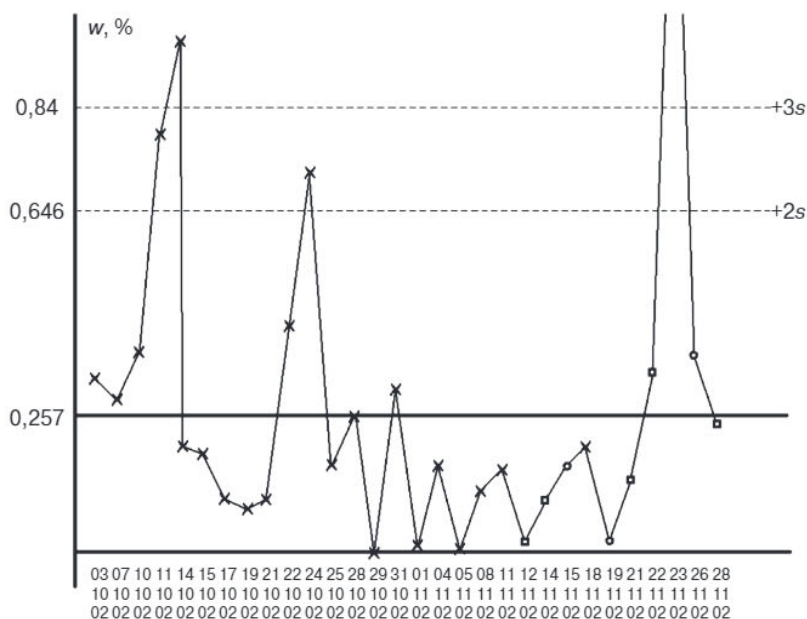


Рис. 4.7. Контрольная карта текущих расхождений. Данные из табл. 4.3

В целом карту *CUSUM* можно считать дополнением к карте Средних значений; карты *CUSUM* несколько лучше нее выявляют дрейфы и небольшие по величине сдвиги среднего значения.

**Контрольная карта текущих расхождений**, показанная для данных из табл. 4.3 на рис. 4.7, аналогична контрольной карте расхождений (размахов) с тем отличием, что вместо расхождений (размахов) величин  $w$  по вертикали откладывают величины текущих расхождений  $w_i = |\bar{X}_{j+1} - \bar{X}_j|$ , то есть разности между средними результатами измерений в следующих одна за другой сериях. Она предназначена для контроля стабильности промежуточной прецизионности (в нашем случае это — промежуточная прецизионность с изменяющимися факторами «время» и «калибровка», характеризуемая стандартным отклонением  $s_{IPT,C}$ ). Конечно, если в каждой аналитической серии выполняется только одно контрольное измерений, его результат и есть среднее. Способы расчета контрольных пределов аналогичны приведенным выше для контрольных карт расхождений (размахов), см. табл. 4.3. Используются те же тревожные и контрольные признаки с той разницей, что вместо расхождения (размаха) рассматривается текущее расхождение.

**Контрольная карта оценок систематической погрешности** представляет собой видоизмененную контрольную карту средних значений, на которой по вертикальной оси вместо средних результатов измерений для подгрупп откладывают оценки систематической погрешности  $\hat{\delta}_j = (\bar{X}_j - \mu)$ , где  $\mu$  — условно истинное или принятое опорное значение измеряемой величины. В качестве горизонтальной оси используют 0, а контрольные пределы  $\pm 1s$ ,  $\pm 2s$  и  $\pm 3s$  наносят в виде горизонтальных линий, аналогично картам средних значений (здесь, как для карт средних значений,  $s \equiv s_{\bar{X}_j}$ ). Такая карта, построенная по данным из табл. 4.3, показана на рис. 4.8.



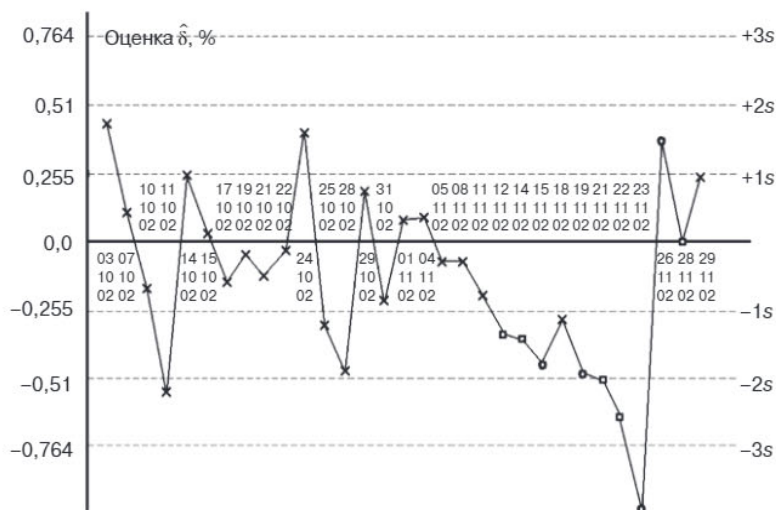


Рис. 4.8. Контрольная карта оценок систематической погрешности. Данные из табл. 4.3

«Тревожный признак» на контрольной карте оценок систематической погрешности — ситуация, когда очередная точка выходит за пределы  $\pm 2s$ . «Контрольный признак»: очередная точка лежит вне пределов  $\pm 3s$ . По этим контрольным картам отслеживают степень отклонения результатов исследований от условно истинного (принятого опорного) значения измеряемой величины.

Конечно, в качестве величины  $\mu$  можно использовать: среднее значение предыдущих результатов измерений в данной лаборатории (то есть  $\mu = \bar{\bar{X}}$ ). Однако это не имеет особого смысла — в этом случае практически получается та же карта средних значений. Поэтому карту оценок систематической погрешности ведут, если в качестве контрольный материал аттестован независимо — когда это стандартный образец; материал, исследованный в авторитетной лаборатории и т. д.

Карты оценок систематической погрешности можно рассматривать как вспомогательные. Причина в следующем. Если величина  $\bar{\bar{X}}$  несущественно отличается от  $\mu$  (то есть систематическая погрешность невелика), то получается карта, близкая к карте Средних значений. Именно такова ситуация в нашем примере (см. табл. 4.3 и рис. 4.8). Если же величины  $\mu$  и  $\bar{\bar{X}}$  существенно различаются, то есть систематическая погрешность велика, то скорее всего так работать нельзя — надо снижать погрешность. В этом случае карту оценок систематической погрешности можно использовать лишь для ориентировки.

**Карты «Экспоненциально взвешенных движущихся средних» (EWMR) [128]** (оригинальное название — геометрические движущиеся средние, GMA) [129]) отличаются от карт *CUSUM* введением весовых коэффициентов, характеризующими вклад предыдущих результатов измерений. Этот вклад тем меньше, чем раньше была проанализирована серия. Карты *EWMR* строят следующим образом.

По окончании периода оценки характеристик рассчитывают общее среднее  $\bar{\bar{X}}$ . Затем рассчитывают значения геометрических движущихся средних  $X^w$  для последующих серий:

$$\begin{aligned}
 X_1^w &= w\bar{X}_1 + (1-w)\bar{\bar{X}}; \\
 X_2^w &= w\bar{X}_2 + (1-w)X_1^w; \\
 &\dots \\
 X_k^w &= w\bar{X}_k + (1-w)X_{k-1}^w; \\
 &\dots
 \end{aligned}
 \tag{4.28}$$

Здесь  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k, \dots$  — средние значения результатов анализа контрольного материала, полученные с момента начала расчета  $EWMR$ ,  $\bar{\bar{X}}$  — общее среднее, полученное к этому моменту в предшествующем периоде.  $w$  — весовой фактор, обычно выбираемый равным 0,20–0,25 [128]. Полученные величины  $X_1^w, X_2^w, \dots, X_k^w, \dots$  наносят на контрольную карту  $EWMR$ , в качестве оси  $X$  используют общее среднее  $\bar{\bar{X}}$ .

Оценка дисперсии величин  $(X_1^w, X_2^w, \dots, X_k^w)$  равна

$$s_{X_k^w}^2 = s^2 \left[ 1 - (1-w)^{2k} \right] \frac{w}{2-w}, \tag{4.29}$$

где  $s \equiv s_{\bar{X}_j}$  — стандартное отклонение средних значений  $\bar{X}_j$ . При больших значениях  $k$  получаем

$$s_{X_k^w}^2 \cong \frac{s^2 w}{2-w}. \tag{4.30}$$

Основным контрольным признаком, используемым при работе с картами  $EWMR$ , является выход очередного результата за пределы  $(\bar{\bar{X}} \pm 3ws)$ .

Как и при использовании  $CUSUM$ , в методе  $EWMR$  в значение последнего результата (по которому ведется контроль) вносят вклад предыдущие. Однако для  $EWMR$  этот вклад тем меньше, чем раньше был получен результат. Можно сказать, что карты  $EWMR$  в этом смысле являются промежуточными между картами средних значений и картами  $CUSUM$ , а то, к какой из них они ближе, определяется выбором весового коэффициента  $w$ .

В графе «Примечания» и в пояснениях к табл. 4.3 даны результаты применения правил Вестгарда, а также правил выявления выхода из-под контроля для других контрольных карт (см. выше), к данным из этой таблицы. Обратим особое внимание на то, что использовавшийся автоанализатор фирмы «Carlo Erba Instr.» является средством измерения, внесенным в Государственный реестр и подлежит обязательной поверке, которая и была проведена летом 2002 г. Формально никакой дополнительный контроль, в том числе и контроль стабильности, не требовался. Однако если бы контроль стабильности не велся, сбои в работе автоанализатора не были бы выявлены (внешних признаков произошедшей поломки не наблюдалось) и не-правильные результаты, возможно, выдавались бы до лета 2003 г.

**«Перекрытие».** Отдельно необходимо рассмотреть ситуацию, когда контроль стабильности уже ведется, и контрольный материал заканчивается (или подходит к концу срок его годности). В этом случае преемственность контроля поддерживают следующим образом.

Незадолго до окончания работы с контрольным материалом (назовем его «выводимым») из-за того, что его количество на исходе или истекает срок хранения, но не позже чем за 20–30 серий до этого момента (в зависимости от продолжительности периода оценки характеристик), начинают анализировать новый контрольный материал с близкими свойствами и концентрацией определяемых веществ (назовем его «вводимым»). На протяжении 20–30 серий измерений выводимый и вводимый контрольные материалы анализируют совместно, причем внутрилабораторный контроль ведут по выводимому контрольному материалу; для вводимого материала — это период оценки характеристик. После этого переходят к контролю по вводимому материалу. Этот способ называют «*перекрыванием*»; он обеспечивает преемственность контроля стабильности.

Конечно, от перекрывания можно отказаться в тех случаях, когда известно, что вводимый контрольный материал точно такой же, как выводимый (например, это стандартный образец водного раствора соли металла той же концентрации). Однако в этом случае мы полностью доверяемся производителю материала.

*Параметры контрольных карт (средние значения, контрольные пределы) рекомендуют уточнять после набора 20–30 очередных серий. При этом часто ограничивают число последних серий, которые берут в расчет, 70–80 сериями, чтобы отсеять совсем старые результаты контроля.*

#### ***О многомерных методах контроля стабильности с использованием контрольных карт***

В последнее время появляется все больше методов КХА, в которых в каждой рутинной пробе и, соответственно, в контрольном материале одновременно измеряется несколько показателей (хромато-масс-спектрометрия, хроматография, атомно-эмиссионный анализ и масс-спектрометрия с использованием индукционно-связанной плазмы и т. д.). При этом при наличии контрольного материала, вообще говоря, возможны два подхода к контролю стабильности измерений:

- результаты определения каждого показателя в контрольном материале, как описано выше, рассматриваются и оцениваются отдельно — одномерный подход;
- все результаты, получаемые в серии для данного контрольного материала, рассматриваются в совокупности — многомерный подход. Кратко рассмотрим его в самом общем виде.

При многомерном статистическом контроле качества<sup>96</sup> в настоящее время используются контрольные карты величин  $T^2$  Хотеллинга [130] и «мер остатков» (*residual measures*)  $Q$  [131]. Эти карты интегрируют многомерные данные и поэтому страдают недостатком наглядности. Эта проблема была решена с появлением «диаграмм вклада» (*contribution plots*) [132], позволяющих определить, какие именно показатели вышли из-под контроля (т. е. какой, собственно, сбой произошел в процессе измерений). Подборка статей по использованию многомерных методов контроля качества опубликована в [133]; см. также [134, 135].

<sup>96</sup> Общеупотребительный английский термин «Multivariate Statistical Process Control (MSPE)» — «многомерный статистический контроль процесса».



Со статистической точки зрения многомерные методы лучше одномерных по следующим причинам:

- меньше вероятность ложной тревоги (ложного отброса контролируемой серии);
- выход из-под контроля фиксируется при маловероятном сочетании результатов измерений нескольких показателей (при этом, возможно, ни на одном из них в отдельности выход из-под контроля не фиксируется).

Однако в химическом анализе, как правило, на результаты определения разных показателей влияют одни и те же факторы, и, следовательно, они коррелируют друг с другом. В таких ситуациях многомерные методы не дают ощутимого выигрыша. Кроме того, если при химическом анализе совместно определяется несколько показателей, при контроле стабильности зачастую можно ограничиться одним-двумя наиболее «капризными» анализатами (подробнее см. ниже, разд. 4.5.3).

#### 4.5.1.4. Контроль стабильности по контрольным картам в отсутствие контрольного материала

Как правило, в большинстве случаев можно найти (приобрести либо приготовить самостоятельно) подходящий контрольный материал. Однако иногда это не удастся, чаще всего из-за нестабильности исследуемых веществ. В этом случае возможны два способа контроля стабильности – с помощью контрольных карт расхождений (размахов) и контрольных карт текущих расхождений, ведущихся по результатам исследований рутинных проб.

**Контрольные карты размахов (расхождений)**, построенные по результатам исследований рутинных проб, аналогичны таковым, построенным на контрольных материалах (см. выше). Отличие в том, что на них откладывается разность между максимальным и минимальным результатами, полученными для какой-либо рутинной пробы.

Пусть в  $j$ -й серии для контроля выбрана некоторая рутинная проба. Обозначим полученные для нее результаты измерений как  $X_{j,1}^{pyt}, X_{j,2}^{pyt}, \dots, X_{j,l}^{pyt}$ . Тогда на контрольную карту будут откладываться разности между максимальной и минимальной из этих величин  $X_{j,max}^{pyt} - X_{j,min}^{pyt}$ . Если, как это чаще всего бывает, для рутинной пробы, выбранной для контроля, в  $j$ -й серии проводится два измерения  $X_{j,1}^{pyt}$  и  $X_{j,2}^{pyt}$ , то на карту откладываются величины  $|X_{j,1}^{pyt} - X_{j,2}^{pyt}|$ .

Основная трудность здесь в том, что пробы, выбранные как контрольные в той или иной серии, должны быть более или менее одинаковыми как по свойствам, так и по значению измеряемой величины. Конечно, это не всегда легко обеспечить, особенно для твердых проб, гомогенность которых не гарантирована. В конкретной серии измерений часто вообще не удастся выбрать подходящую пробу (например, в экологической лаборатории в какой-то серии вообще может не оказаться проб с достоверно обнаруживаемой концентрацией анализата).

Требование близости значений определяемого показателя (концентрации анализата) для разных рутинных проб, используемых для контроля, в основном связано с тем, что повторяемость (сходимость) зависит от измеряемой величины. Правда, влияние этого фактора можно уменьшить (но не устранить), нанося на контрольную

карту величины размахов (расхождений), отнесенные к приписанным методике значениям стандартного отклонения повторяемости (сходимости) [117].

Даже если такой контроль удастся наладить, он оказывается неэффективным — контролируется стабильность лишь одного из компонентов прецизионности — повторяемости (сходимости), вклад которой в общую погрешность невелик. Поэтому всегда необходимо стремиться к контролю стабильности по контрольному материалу.

**Контрольные карты текущих расхождений** также могут вестись по результатам исследований рутинных проб следующим образом.

В первой серии выбирается и исследуется одна из рутинных проб и по результатам для нее рассчитывается среднее значение. Часть этой пробы сохраняют до второй серии, где вновь исследуют, после чего для нее рассчитывают разность между средними значениями в первой и второй сериях и наносят абсолютную величину этой разности на контрольную карту.

Кроме пробы из первой серии, во второй серии выбирают и исследуют другую рутинную пробу, рассчитывают для нее среднее значение и часть ее сохраняют до третьей серии. Исследовав эту пробу в третьей серии и рассчитав среднее, вычитают из него предыдущее среднее и абсолютную величину разности наносят на контрольную карту.

Кроме пробы из второй серии, в третьей серии выбирают и исследуют третью рутинную пробу, рассчитывают для нее среднее значение и часть ее сохраняют до четвертой серии и так далее.

Все это можно записать следующим образом. Обозначим среднее значение результатов исследования рутинной пробы, выбранной для контроля из проб  $j$ -й серии, в  $j$ -й и  $j+1$  сериях, как  $\bar{X}_{j(j)}^{\text{рут}}$  и  $\bar{X}_{j(j+1)}^{\text{рут}}$ . На контрольную карту текущих расхождений откладываются абсолютные значения разностей этих величин  $|\bar{X}_{j(j)}^{\text{рут}} - \bar{X}_{j(j+1)}^{\text{рут}}|$ . В остальном контрольная карта текущих расхождений по результатам исследования рутинных проб ведется также, как аналогичная карта, построенная для контрольного материала.

Контрольные карты текущих расхождений по рутинным пробам ведут крайне редко. Дело в том, что отсутствие контрольного материала, как правило, связано с нестабильностью объекта исследований. А для таких объектов почти никогда нельзя гарантировать неизменность определяемого показателя в рутинных пробах до следующей серии измерений. Кроме того, организацию такого контроля усложняют все те факторы, которые мешают вести контроль повторяемости (стабильности) по рутинным пробам (см. выше).

#### 4.5.2. Другие способы внутрилабораторного контроля

Эти способы, большинство которых изложено в [117], заключаются в том, что тем или иным способом (анализ стандартных образцов, данные об анализе рутинных проб, введение добавок и т. д.) получается выборка, отражающая метрологические характеристики измерений за какой-то период времени. Получая оценки характеристик этой выборки (среднее значение, стандартное отклонение) и сравнивая их с приписанными значениями, судят о качестве измерений за оцениваемый период.



Общим недостатком таких способов по сравнению с контролем стабильности по контрольным картам является неиспользование сведений о ходе процесса измерений в динамике и игнорирование при их применении почти всегда имеющегося «запаса качества». Именно поэтому они не рекомендованы международным стандартом [27].

### *Простейший оперативный контроль<sup>97</sup>*

Простейший оперативный контроль предполагает точечную оценку качества и имеет несколько разновидностей.

*Простейший оперативный контроль с использованием аттестованных материалов* состоит в том, что в какой-либо серии выполняется несколько измерений для какого-либо аттестованного материала (обычно стандартного образца, СО). Часто полагают, что число измерений при этом должно быть таким же, как для рутинных проб. Это ниоткуда не следует, и для большей надежности контроля всегда лучше выполнить несколько (не менее 2–3) контрольных измерений. Периодичность такого контроля выбирают, исходя из опыта — чем стабильнее методика, тем реже можно проводить контроль. Предположим, что мы исследовали СО с аттестованным значением  $X_a$  и погрешностью этого значения  $\Delta X_a$ , выполнили  $N$  измерений в одной серии и получили результаты  $X_1, X_2, \dots, X_p, \dots, X_N$ . Тогда возможны два случая.

Если погрешностью  $\Delta X_a$  можно пренебречь, систематическую погрешность можно считать не выявленной при выполнении условия

$$|\bar{X} - X_a| < z\sigma_R, \quad (4.31)$$

где  $\bar{X}$  — среднее величин  $X_p$ ,  $\sigma_R$  — стандартное отклонение воспроизводимости, соответствующее значению определяемой величины (во многих случаях указанное в методике),  $z$  — выбранный квантиль нормального распределения (чаще всего его принимают равным 2, что соответствует доверительной вероятности  $\sim 0,95$  для двухстороннего критерия)<sup>98</sup>.

В противном случае необходимо проверить гипотезу о равенстве нулю систематической погрешности. Проще всего сделать это, посмотрев, пересекаются ли интервалы  $(X_a \pm \Delta X_a)$  и  $(\bar{X} \pm z\sigma_R)$ . Если да, то систематическая погрешность не выявлена. Конечно, этот вариант хуже, однако не всегда можно найти подходящий стандартный образец.

Если значение  $\sigma_R$  в методике отсутствует, для корректного использования данного способа контроля его необходимо оценить. Если в методике приведен предел воспроизводимости  $R$ , значение  $\sigma_R$  можно найти как  $\sigma_R = R/2,77$ . Возможно также использование стандартной и расширенной неопределенности (см. разд. 3.5.2).

<sup>97</sup> В [117] подобные процедуры называются «оперативным контролем» — термин не слишком удачный, поскольку контроль стабильности с помощью контрольных карт по меньшей мере столь же оперативен.

<sup>98</sup> Часто правую часть формулы (ВК20) делят на 1,2 (добавляют множитель 0,84), а также требуют повторения измерения в случае неудовлетворительного результата для снижения вероятности ошибки второго рода.



Если в ходе проверки достоверно выявлена систематическая погрешность, необходимо принять меры к ее снижению. Заметим, что в случае, когда ведется и контроль стабильности по контрольным картам, и (независимо) простейший оперативный контроль, и результаты противоречат друг другу, мы оказываемся в непростой ситуации — приходится неформально разбираться в причинах расхождения.

**Простейший оперативный контроль повторяемости (сходимости) по размахам** состоит в том, что для какой-либо пробы (рутинной, СО и т. д.) расхождение (размах) сравниваются с пределом повторяемости (сходимости)  $r$  методики  $s_r$  при двух измерениях ( $n = 2$ ) либо с величиной  $CR_{0,95}(n)$  при  $n > 2$  (см. разд. 4.3.3 и табл. 4.1). Превышение полученной величины над нормативной свидетельствует о неприемлемой повторяемости (сходимости). Такая проверка фактически совпадает с тем, что делается при проверке приемлемости результатов на рутинных пробах (см. разд. 4.3.3). И, конечно, такой контроль не нужен, если ведется контроль повторяемости (сходимости) по контрольным картам размахов.

Несколько разновидностей **простейшего оперативного контроля, связанных с введением добавок и разведением пробы**, также рекомендованы в [117]. Однако, прежде чем применять эти методы, необходимо понять, когда их можно применять и что они дают.

Способы контроля, связанные с введением добавок, можно применять при двух условиях:

- необходимо иметь возможность ввести добавку в той же форме, в которой определяемое вещество содержится в пробе;
- надо уметь заведомо равномерно распределять добавку по объему пробы.

Для твердых проб гарантировать выполнение этих условий в большинстве случаев невозможно, поэтому введение добавок, если и используется, то в основном для жидких проб. То же касается и способов, сопряженных с разбавлением пробы.

**Введение добавок** при контроле качества химического анализа используют для проверки правильности методики (отсутствия достоверно выявленных систематических погрешностей) в отсутствие стандартных образцов. В одну из рутинных проб вносят одну или несколько добавок определяемого вещества, измеряя его концентрацию до и после введения добавок. При этом исходная проба может не содержать либо содержать определяемое вещество.

В литературе используются два термина: «метод добавок» и «метод «введено-найден»».

Термин «введено-найден» (соответствует английскому «recovery») используется в случаях, когда добавка вводится для доказательства правильности методики. Термин «метод добавок» интерпретируется по-разному. В аналитической химии он полностью соответствует английскому «standard additions method» и используется в случаях нахождения концентрации аналита с помощью добавок. В [117] термин «метод добавок» используется вместо «введено-найден» для способа проверки правильности методики. Это противоречит общепринятой терминологии и, можно надеяться, оно будет исправлено. Мы будем использовать оборот «введено-найден».

Рассмотрим случай введения одной добавки (обычно ее выбирают близкой к концентрации определяемого вещества в пробе). Обозначим результаты измерений без добавки как  $X_1, X_2, \dots, X_N$  (среднее значение  $\bar{X}$ , стандартное отклонение  $s_x$ ,

число измерений  $N_1$ ), а с добавкой как  $Y_1, Y_2, \dots, Y_{N_2}$  (среднее значение  $\bar{Y}$ , стандартное отклонение  $s_Y$ , число измерений  $N_2$ ). Тогда величина  $t = \frac{|\bar{Y} - \bar{X} - M|}{s} \sqrt{\frac{N_1 N_2}{N_1 + N_2}}$

(где  $M$  – точно известная добавка,  $s = \sqrt{\frac{s_X^2 (N_1 - 1) + s_Y^2 (N_2 - 1)}{N_1 + N_2 - 2}}$ ) подчиняется  $t$ -распре-

делению Стьюдента<sup>99</sup>. Найденное значение  $t$  сравнивают с процентной точкой  $t$ -распределения  $t_{\text{крит}}(\bar{P}, f)$ , соответствующей выбранной доверительной вероятности  $\bar{P}$  (обычно 0,95) и числу степеней свободы  $f = N_1 + N_2 - 2$ . Если  $t > t_{\text{крит}}(\bar{P}, f)$ , то гипотезу об отсутствии систематической погрешности отвергают. В этом случае необходимо найти причину выявленной погрешности и устранить ее.

Возможно также введение разных добавок в несколько одинаковых проб с совместной обработкой полученных результатов, как это делается при аттестации методик КХА (см. разд. 3.3.4).

Метод «введено-найден», помимо вышеописанных ограничений, имеет следующий недостаток. Поскольку дисперсия  $s^2$  разности  $(\bar{X} - \bar{Y})$  есть сумма дисперсий  $\bar{X}$  и  $\bar{Y}$ , погрешность, возникающая при расчетах и рассчитываемая как погрешность разности двух больших величин, весьма велика. При этом мы «зажаты» с двух сторон: увеличение добавки снижает вклад матричных эффектов, а уменьшение ее – к снижению достоверности выявления эффекта. Именно поэтому при введении одной добавки ее выбирают так, чтобы она добавляла концентрацию, близкую к имеющейся в пробе. Для повышения мощности метода «введено-найден» целесообразно выполнять несколько измерений для проб с добавкой и без. Желательно, чтобы  $N_1$  и  $N_2$  лежали бы в диапазоне 4–8, но так почти никогда не делают из-за высокой трудоемкости.

В любом случае регулярное применение метода «введено-найден» при внутрिलाбораторном контроле не имеет особого смысла. Дело в том, что систематические погрешности, связанные с матричным эффектом, практически никогда не возникают вдруг – для этого должны произойти существенные изменения в методике (изменение процедуры анализа, введение новых реактивов и т. д.). Если этого не было, метод «введено-найден» практически никогда не выявляет погрешности. А «обычные» пропорциональные систематические погрешности (например, связанные с изменением чувствительности, связанным с градуировкой) этим методом не выявляются в принципе.

**Метод разведения пробы.** Какую-либо пробу делят на две части, одну из них просто анализируют, а вторую исследуют после разведения в  $\eta$  раз, после чего анализируют каждую из них несколько раз в одной серии. Далее проверяют гипотезу о равенстве математических ожиданий результатов измерений для исходной пробы и результатов измерений для разбавленной, умноженных на  $\eta$ .

Обозначим результаты измерений для исходной пробы как  $X_1, X_2, \dots, X_{N_1}$  (среднее значение  $\bar{X}$ , стандартное отклонение  $s_X$ , число измерений  $N_1$ ), а для разбавленной в  $\eta$  раз как  $Y_1, Y_2, \dots, Y_{N_2}$  (среднее значение  $\bar{Y}$ , стандартное отклонение  $s_Y$ , число измерений  $N_2$ ). Умножим все величины  $Y$  на  $\eta$ . Тогда величина

<sup>99</sup> Формула справедлива, если  $s_X$  и  $s_Y$  не различаются достоверно (см. разд. 2.5.3).



$$t = \frac{|\eta \bar{Y} - \bar{X}|}{s} \sqrt{\frac{N_1 N_2}{N_1 + N_2}}, \text{ где } s = \sqrt{\frac{s_X^2 (N_1 - 1) + \eta s_Y^2 (N_2 - 1)}{N_1 + N_2 - 2}}, \text{ подчиняется } t\text{-распределению}$$

Стьюдента<sup>100</sup>. Найденную величину  $t$  сравнивают с процентной точкой  $t$ -распределения  $t_{\text{крит}}(\bar{P}, f)$ , соответствующей выбранной доверительной вероятности  $\bar{P}$  (обычно 0,95) и числу степеней свободы  $f = N_1 + N_2 - 2$ . Если  $t > t_{\text{крит}}(\bar{P}, f)$ , то гипотезу об отсутствии систематической погрешности отвергают. В этом случае необходимо найти причину появления выявленной (систематической) погрешности и устранить ее.

Метод разведения, также как метод «введено-найдено», пригоден преимущественно для водных растворов.

*Метод разведения с добавкой* является комбинацией двух предыдущих методов (пробу сначала разводят, а затем вводят добавку).

В целом можно сказать, что использование простейшего оперативного контроля при систематически выполняемых исследованиях малоэффективно. При этом способы, связанные с введением добавки и разбавлением пробы, весьма трудоемки. Использование их имеет смысл при аттестации методик, а также в случае редко выполняемых измерений.

#### «Статистический»<sup>101</sup> контроль качества

«Статистический» контроль качества (по альтернативному признаку) имеет множество разновидностей. Все они основаны на идее выборочного контроля, подобного выборочному контролю качества изделий в промышленности, общая схема которого такова:

- определяется партия изделий (например, гайки, отштампованные в течение одной смены);
- по статистическим таблицам определяется объем выборки изделий для контроля с учетом допустимой доли брака («приемлемый уровень качества»);
- из партии извлекается (случайным образом) нужное количество изделий для выборочного контроля;
- каждое изделие из выборки тестируется и относится либо к дефектным, либо к качественным;
- количество дефектных (бракованных) изделий сравнивается с табулированным критическим значением, зависящим от допустимой доли брака, и в случае превышения вся партия бракуется.

Возможны более сложные варианты этой схемы (для уменьшения объема работы контроль может проводиться в несколько этапов), но общий принцип сохраняется.

Для лабораторного анализа при определении партии проб возможны две ситуации.

<sup>100</sup> Формула справедлива, если величины  $s_X$  и  $s_Y$  не различаются достоверно (см. разд. 2.5.3).

<sup>101</sup> Здесь термин «статистический» использован в соответствии с [117]. Следует отличать «статистический» контроль в узком смысле, описанный в этом разделе, от статистического контроля вообще (см. выше, разд. 4.3).



1. Имеются партии проб, анализируемых совместно и заданные их происхождением (партия проб, взятых с определенной территории или месторождения в геологии, пробы, являющиеся выборкой из какой-либо популяции при популяционных исследованиях в биологии и медицине [63] и т. д.).
2. «Партия» проб определяется искусственно, например, как все пробы, которые будут проанализированы в течение определенного периода времени: месяца, квартала и т. д.

В первом случае, вообще говоря, контролируется не качество анализа как такового, а качество исследования конкретной партии проб (одновременно могут анализироваться как несколько партий проб, так и отдельные пробы, не входящие в какую-либо партию). Второй вариант, собственно, и предназначен для контроля качества анализа.

Роль выбранных из партии для тестирования изделий при химическом анализе играют различные «контрольные процедуры»:

- повторный анализ какой-либо рутинной пробы в той же серии, в другой серии либо в другой лаборатории;
- исследование стандартных образцов;
- оценка качества методом «введено-найдено» и т. д.

Другими словами, одновременно с анализом некоторого числа рутинных проб выполняют «контрольные процедуры», число которых выбирается так же, как объем выборки изделий (см. выше). Для каждой «контрольной процедуры» устанавливается критерий удовлетворительности (брак — не брак); в целом это похоже на простейший оперативный контроль. Если доля брака превышает допустимую, партия бракуется, то есть результаты анализа всех проб, входящих в партию, считаются неудовлетворительными.

В первой ситуации (когда целью является только исследование партий проб, называемых также выборками из популяций [63]), результаты «статистического» контроля является естественным дополнением к результатам анализа, подтверждающим его достоверность. Однако для контроля качества анализа в целом, вне привязки к партиям проб, «статистический» контроль малоэффективен из-за отсутствия оперативности. Действительно, результаты статистической оценки качества анализа рутинных проб появляются только в конце длительного периода времени (обычно через 2—3 месяца после начала периода оценки), когда большинство результатов анализа уже выдано заказчикам и использовано ими. Трудно себе представить отзыв этих результатов и повторный анализ проб, исследованных за контролируемый период — и, конечно, никто не станет браковать большие массивы полученных результатов.

«Статистический» контроль в отсутствие партий проб можно рассматривать как дополнение к процедуре простейшего оперативного контроля, включающее планирование такого контроля и последующую статистическую обработку.

**Использование результатов проверки приемлемости для контроля повторяемости (сходимости) анализа в целом.** Алгоритмы проверки приемлемости (см. разд. 4.4) предполагают обязательную оценку расхождения (размаха) результатов измерений для каждой рутинной пробы. Помимо использования результатов такой проверки в алгоритме получения окончательно приводимого результата измерений, по ним можно судить о повторяемости (сходимости) процесса измерений в целом.

Исходя из основной модели процесса измерений (4.10) и общих принципов организации внутрилабораторного контроля, естественно считать неудовлетворительной ту или иную серию измерений, если число рутинных проб, для которых размах превысил величину  $CR_{0,95}(N)$ , где  $N$  — число измерений для одной рутинной пробы, слишком велико. Оценить это можно следующим образом.

Предположим самый плохой из допустимых случаев, когда стандартное отклонение повторяемости (сходимости)  $s_r$  равно нормативу  $\sigma_r$ . Пусть число рутинных проб в серии равно  $U$ , и в ней неудовлетворительными по размаху оказались результаты исследования  $V$  проб. Тогда вероятность такого события можно рассчитать по формуле Бернулли

$$P_U(V) \cong C_U^V (0,05)^V (0,95)^{U-V} = \frac{U!}{V!(U-V)!} (0,05)^V (0,95)^{U-V}. \quad (4.32)$$

Здесь  $C_U^V$  — число сочетаний из  $U$  по  $V$ , 0,05 и 0,95 — округленные вероятности появления неудовлетворительного и удовлетворительного результата соответственно. Данную серию можно считать неудовлетворительной, если величина  $PU(V)$  меньше принятой вероятности ошибки первого рода  $P$ . Чаще всего принимают  $P = 0,01$  (доверительная вероятность  $\bar{P} = 0,99$ ).

**Пример.** В конкретной серии исследовали 15 рутинных проб, и для двух из них размах превышал допустимый. Вероятность такого события равна  $0,0135 > 0,01$ , то есть серия не признается неудовлетворительной. В то же время такой результат должен вызывать тревогу, поскольку вероятность такого события меньше 0,05.

Отметим, что в отличие от контроля повторяемости (сходимости) с помощью контрольных материалов при использовании для контроля рутинных проб могут возникать проблемы, связанные с их разной неоднородностью. Оценки повторяемости (сходимости) на неоднородном образце будут завышенными, и возникнет ложная тревога.

#### 4.5.3. Выбор оптимальных способов ведения внутрилабораторного контроля в конкретных ситуациях

Проверку приемлемости необходимо проводить всегда, когда это возможно, то есть при двух или более измерениях для рутинных проб. Это позволяет с высокой вероятностью избежать грубых промахов при анализе этих проб. Кроме того, результаты проверки приемлемости можно использовать для контроля качества аналитического процесса в целом (см. выше).

В то же время, как видно из предыдущего текста, существует достаточно много вариантов ведения внутрилабораторного контроля для каждой методики в конкретной ситуации, и выбрать оптимальный в каждом случае бывает непросто. Можно предложить некоторый алгоритм, помогающий сделать этот выбор с учетом эффективности каждого из видов контроля. Предположим, что для некоторой методики внутрилабораторный контроль ранее не велся и мы хотим спланировать его «с нуля».

Прежде всего необходимо задаться вопросом, насколько часто будет выполняться данный вид анализа. Если не реже, чем 1–2 раза в неделю, оптимальным



является контроль стабильности с помощью контрольных карт. Если не так часто («редко выполняемый анализ»), то возможно три основных варианта:

- не организовывать контроль вообще, а передавать пробы для данного анализа в другую лабораторию, в которой такие исследования выполняются часто и налажен контроль качества (субподряд). Это вполне логично — не надо будет каждый раз мучиться с обеспечением и доказательством качества измерений, да и с экономической точки зрения выгоднее. К сожалению, передача анализа на субподряд не всегда возможна либо из-за отсутствия подходящей лаборатории, либо из-за противодействия со стороны администрации, считающей финансирование субподрядных работ такого рода нерациональным;
- выполнять данный вид исследований чаще, чем это на самом деле необходимо (например, один раз в неделю), контролируя при этом стабильность по контрольным картам. В этом случае лаборатория всегда будет готова выполнить редкое измерение с требуемым качеством. Такой вариант приемлем только в случае простых и дешевых измерений (например, измерение радиоактивности).
- организовывать при необходимости определения редкого показателя (анализа редкого объекта) серию специального дизайна. При наличии аттестованного материала (например, адекватного стандартного образца) можно выполнить несколько (желательно не менее 8) измерений, что позволяет оценить и правильность, и сходимость в этой серии. Одновременно (желательно попеременно с аттестованным материалом) анализируются рутинные пробы. Удовлетворительные как по сходимости, так и по правильности результаты исследования аттестованного материала более или менее гарантируют качество исследования рутинных проб. По сути эта процедура похожа на простейший оперативный контроль с использованием стандартного образца, но с увеличенным числом измерений.

Если интенсивность выполнения данного вида анализа достаточно велика (не реже 1–2 раз в неделю), то необходимо определить, доступны ли в достаточном количестве контрольные материалы — стабильные в течение длительного времени и однородные вещества, похожие на исследуемые в лаборатории рутинные пробы. Желательно, чтобы срок годности их был бы не менее года, а количество позволяло бы вести контроль по избранному плану см. разд. 4.3.1.1) в течение этого времени с учетом интенсивности работы.

Если такие материалы недоступны, контроль ведут по контрольным картам размахов и, крайне редко, текущих расхождений, которые строят по результатам исследования рутинных проб (см. разд. 4.5.1.4). При наличии стандартных образцов в дополнение к этому периодически (обычно не чаще 2–4 раз в месяц) проводят простейший оперативный контроль с их использованием. Могут использоваться и другие способы простейшего оперативного контроля.

В большинстве случаев контрольные материалы имеются (могут быть приобретены либо изготовлены в лаборатории). В этом случае ведется полноценный контроль стабильности с их использованием, подробно описанный выше. Некоторое различие возникает между ситуацией, когда контрольный материал аттестован (например, если это стандартный образец или аттестованная смесь) и если он не



аттестован. Оно состоит в том, что для аттестованных материалов может вестись контрольная карта оценок систематической погрешности, а для не аттестованных — нет.

В целом можно сказать, что внутрилабораторный контроль качества стал мощным инструментом обеспечения качества рутинного химического анализа. Как показывает практика, вышеописанные алгоритмы и подходы позволяют эффективно выявлять и во многих случаях предупреждать выход аналитического процесса из статистически контролируемого состояния.

Вместе с тем необходимо четко понимать, что внутрилабораторный контроль есть лишь один из способов обеспечения качества и возможности его ограничены. Это определяется рядом обстоятельств.

Во-первых, внутрилабораторный контроль в описанном выше виде не охватывает процедуру отбора пробы, которая при анализе неоднородных объектов вносит существенный, а часто и определяющий вклад в общую погрешность. Стадия пробоотбора нуждается в особых способах планирования и контроля (см. разд. 4.2).

Во-вторых, удовлетворительные результаты внутрилабораторного контроля не являются абсолютной гарантией такого же качества анализа исследуемых рутинных проб. Различия в качестве анализа контрольных материалов и рутинных проб могут возникать как из-за различия в составе их матрицы (неадекватность), так и в силу других, в том числе случайных, причин.

В-третьих, существует субъективный фактор — выполняющий анализ персонал часто по разным причинам заинтересован в «улучшении» (подгонке) результатов внутрилабораторного контроля. Иногда это проявляется в том, что результаты контроля стабильности «группируются» вблизи контрольных пределов ( $\bar{X} \pm 2s$ ). Все это, безусловно, снижает эффективность контроля.

Тем не менее правильно организованный внутрилабораторный контроль является важнейшим способом обеспечения качества химического анализа и необходим при выполнении рутинных исследований. Рассмотренные выше подходы используются для контроля качества как химического анализа, так и других видов лабораторных исследований.

*Ведение внутрилабораторного контроля вообще и контроля стабильности, в особенности вручную, весьма трудоемко, что часто является препятствием для его внедрения в лабораториях. То же касается и процедур обеспечения качества в целом. Проблема решается с использованием компьютеров (подробнее см. разд. 4.7).*

## 4.6. Внешняя оценка качества результатов количественного химического анализа

Разновидностью межлабораторных экспериментов (см. разд. 3.6) являются программы внешней оценки качества результатов химического анализа (Межлабораторные сравнительные испытания (МСИ), англ. Proficiency Testing, External Quality Assessment) — ниже «Программы». Они отличаются от других видов межлабораторного эксперимента задачей — их основной целью является проверка соответствия качества анализа в каждой из лабораторий-участниц эксперимента установленным

требованиям [136]. Дизайн и способы обработки данных в таких программах обычно упрощены: как правило, в лаборатории рассылается один или два однотипных образца и используется либо простейший дизайн (одно измерение для образца), либо каждой лаборатории предлагают выполнить два измерения для каждого образца в одной аналитической серии. Постоянно действующие программы внешней оценки качества фактически состоят из периодически проводимых межлабораторных экспериментов, каждый из которых обычно называется циклом оценки качества.

Организация, осуществляющая Программу, называется «Провайдер».

Важнейшую роль в организации любой Программы внешней оценки качества как, впрочем, и при организации других видов межлабораторных экспериментов, играют их координаторы. Ответственность, которую они несут, очень велика — оценки, получаемые лабораториями, могут сыграть огромную роль в судьбе этих лабораторий. Поэтому к координаторам Программ предъявляются высокие требования — как профессиональные, так и в чисто человеческом плане. Координатор должен обладать соответствующей квалификацией и опытом работы в сфере организации и проведения межлабораторных экспериментов (или тесно сотрудничать с лицами, имеющими такую квалификацию и опыт), работать без какого-либо давления извне, быть компетентным как в области тематики исследований, так и в обработке данных.

При необходимости координатор создает рабочую группу из специалистов (экспертов), которые участвуют в подготовке протокола Программы и решении вопросов, возникающих в ходе ее реализации.

До начала реализации Программы составляется протокол, в котором должна содержаться следующая информация:

- название и адрес организации, осуществляющей Программу;
- фамилия и адрес координатора и других специалистов, участвующих в выполнении Программы;
- цели и задачи Программы;
- критерии выбора лабораторий-участников Программы;
- информация о лабораториях, выполняющих Программу или какие-то ее фрагменты (например, о подготовке контрольных материалов или их проверке на однородность и стабильность);
- информация об исследуемых веществах, определяемых компонентах и методиках анализа с кратким обоснованием их выбора;
- сведения о контрольных материалах (способы их приготовления, проверки и транспортировки);
- график получения лабораториями-участниками Программы необходимой информации и контрольных материалов, а также сроки выполнения анализов и оценки их результатов (для постоянно действующих программ — периодичность оценки, например число циклов оценки в год);
- схема анализа контрольных проб в лабораториях-участниках Программы (например, количество измерений для каждой контрольной пробы);
- описание, хотя бы в общих чертах, методов статистической обработки данных, включая способы выявления «выбросов» и нахождения аттестованных значений;



- перечень сведений, направляемых лабораториями-участницами координатору;
- критерии Программы (критерии, по которым результаты признаются неудовлетворительными);
- перечень документов, направляемых в лаборатории-участницы Программы (оценки, свидетельства об участии в Программе и т. д.) с указанием сроков или периодичности их рассылки;
- описание степени конфиденциальности полученных в Программе результатов и т. д.

При внешней оценке качества чаще всего лабораториям передают не аттестованный контрольный материал и устанавливают среднее содержание определенных веществ по результатам эксперимента (см. разд. 3.1.4). При этом используются обычные приемы для получения объективных сведений о качестве анализа в лабораториях (см. разд. 3.6.2). Конечно, результаты программ внешней оценки качества дают, кроме собственно оценок качества работы лабораторий, всю информацию, какую получают и в других видах межлабораторного эксперимента.

По результатам внешней оценки качества каждой лаборатории дается оценка. Критерии оценки формируются в каждой программе и заносятся в протокол Программы. Их выбор может вытекать либо из нормативных документов, либо просто из возможностей большинства лабораторий.

В простейшем случае в нормативном документе непосредственно указывается критерий, которому должна соответствовать каждая лаборатория. В качестве такого критерия может быть: при одном измерении в лаборатории — максимально допустимая (по абсолютной величине) разность между результатом измерения и аттестованным (опорным) значением; при нескольких измерениях в лаборатории — также максимально допустимая (по абсолютной величине) погрешность, рассчитанная как разность между средним значением результатов, полученных в лаборатории, и аттестованным значением. При нескольких измерениях одного материала для оценки повторяемости (сходимости) используется максимально допустимый размах результатов, полученных в лаборатории.

Возможны несколько подходов к оценке качества работы лаборатории.

Часто в документе на методику количественного химического анализа указывается максимально допустимое расхождение результатов единичных измерений, полученных в двух лабораториях, например: «Два результата, полученные в разных лабораториях, признаются достоверными (с доверительной вероятностью 0,95), если расхождение между ними не превышает  $D$ »; при этом указывается конкретное значение предела воспроизводимости  $D$  в единицах концентрации. В более «свежих» методиках вместо  $D$  используется обозначение  $R$ . Также вместо предела воспроизводимости может приводиться стандартное отклонение воспроизводимости  $\sigma_R = R/2,77$  (см. разд. 3.3.1) — в этом случае расчет величины  $R$  также не составляет труда. Тогда удовлетворительными чаще всего считают результаты, лежащие в интервале {аттестованное значение  $\pm 0,5R$ }. Конечно, это возможно, если все лаборатории-участники Программы работают по аттестованным методикам.

При отсутствии нормативных документов, содержащих фиксированные критерии, возможны два подхода.



- критерии устанавливаются решением рабочей группы с учетом как назначения результатов исследований, проводимых лабораториями, так и реальных возможностей этих лабораторий.
- критерии вырабатываются в зависимости от полученных в данном цикле результатов. Чаще всего заранее выделяются те лаборатории, чьи результаты анализа считаются удовлетворительными. Это делается следующим образом.

Рассчитывают (после отброса выпадающих результатов) статистические показатели – среднее значение  $\bar{X}$  и стандартное отклонение  $s$  результатов измерения концентрации какого-либо вещества в контрольном материале, полученных в очередном цикле Программы. В случае существенных отклонений распределения результатов от нормального его предварительно нормализуют. Удовлетворительными считают результаты, лежащие в интервале  $(\bar{X} \pm ks)$ , причем коэффициент  $k$  выбирают в зависимости от жесткости требований, предъявляемых к лабораториям – обычно в пределах 1,6–2,5.

Практика показывает, что варьирование от цикла к циклу критериев оказывается полезным. Лаборатории, получившие неудовлетворительные оценки, стараются улучшить свою работу, и со временем общий разброс результатов уменьшается.

При оценке качества работы лабораторий часто используют так называемые  $Z$ -индексы ( $Z$ -scores), рассчитываемые для каждой  $i$ -лаборатории по каждому определяемому веществу

$$Z = \frac{|X_i - X_{\text{атт}}|}{s}, \quad (4.33)$$

где  $Z$  –  $Z$ -индекс,  $X_i$  – результат, полученный в  $i$ -й лаборатории;  $X_{\text{атт}}$  – аттестованное значение для контрольного материала (чаще всего среднее значение, полученное в данном цикле оценки);  $s$  – «оценка/мера изменчивости, которая выбирается так, чтобы она удовлетворяла требованиям Программы» [136]. В большинстве случаев (но не всегда) величину  $s$  принимают равной стандартному отклонению результатов, полученных в данном цикле Программы во всех лабораториях (после отброса выпадающих значений).

Часто  $Z$ -отметки округляют до целого числа. Лаборатории ранжируют по возрастанию  $Z$ -отметок и качество работы лабораторий с  $Z$ -отметками более 2 считают неудовлетворительным (такой подход практически не отличается от предыдущего при  $k = 2$ ). Во многих случаях явно или неявно предполагают также, что качество анализа в лабораториях с малыми  $Z$ -отметками выше, чем в лабораториях с большими отметками<sup>102</sup>.

В качестве примера использования разных способов оценки качества в Приложении 10 дано краткое описание одной из систем внешней оценки качества (химический анализ в медицине); более детальное описание приведено в [137].

<sup>102</sup> На самом деле это не всегда так. Все результаты, попадающие в доверительный интервал для среднего значения, неразличимы (см. разд. 2.5.3.1). Кроме того, при расчете  $Z$ -отметок результаты, полученные в лабораториях, рассматриваются как случайная выборка, и не удивительно, если в следующем цикле лаборатории получат совсем другие  $Z$ -отметки.

Результаты внешней оценки качества играют важную роль при аккредитации лабораторий (см. разд. 3.7).

## **4.7. Компьютеризация при построении системы обеспечения качества аналитических лабораторий [138]**

Компьютеризация наряду с разработкой новых и совершенствованием имеющихся методов анализа является магистральным путем развития аналитических и других лабораторий. Однако конкретные пути использования компьютеров в различных лабораториях могут сильно отличаться, и вопрос выбора того или иного из них становится весьма актуальным.

### **4.7.1. Для каких целей применяются компьютерные программы в лабораториях**

Исходя из природы лабораторных исследований, а также с учетом современных требований к обеспечению их качества, сформулированных в [73, 27], можно выделить следующие основные виды лабораторной деятельности, которые могут подлежать компьютеризации:

1. Управление аналитическими приборами и получение аналитического сигнала;
2. Градуировка в методиках измерений;
3. Расчет результатов измерений;
4. Оценка метрологических характеристик методик при их постановке в лаборатории («валидация» методик);
5. Проверка приемлемости результатов;
6. Контроль стабильности результатов;
7. Другие способы внутрилабораторного контроля;
8. Регистрация поступающих в лабораторию проб и отслеживание их движения внутри лаборатории;
9. Подготовка протоколов и их передача заказчикам исследований.
10. Учет и хранение документов (включая методики выполнения измерений), изменений и дополнений к ним, архивных копий;
11. Хранение и использование перечней определяемых показателей и объектов исследований;
12. Ведение списков заказчиков и субподрядчиков;
13. Учет реактивов, материалов и стандартных образцов;
14. Учет и контроль состояния оборудования (средств измерений, испытательного и вспомогательного оборудования); Регистрация претензий и предпринятых корректирующих действий;
15. Учет сотрудников и слежение за своевременностью повышения ими квалификации;
16. Планирование и контроль результатов внутреннего аудита и анализа со стороны руководства;



17. Печать сведений о лаборатории для использования как самой лабораторией, так и внешними организациями (например, органами по аккредитации);
18. Хранение информации.

Этот список не является полным (в него не включены, например, планирование в тех или иных формах, взаимодействие с органами по аккредитации и исследований, экономические расчеты и т. д.), однако он вполне достаточен для наших целей.

Рассмотрим вышеперечисленные задачи по пунктам с позиций возможности и необходимости их компьютеризации.

*Управление аналитическими приборами и получение аналитического сигнала (1)* — традиционная сфера применения вычислительной техники. Компьютеризация приборов, включающая в себя управление аналитическими приборами, регистрацию аналитического сигнала (а также, во многих случаях, расчет результатов измерений) получила в последние годы огромное развитие. Используются как специализированные чипы и компьютеры, так и персональные компьютеры. В большинстве случаев эти устройства являются неотъемлемой частью того или иного аналитического прибора или устройства. Мы не будем здесь рассматривать их детальнее, поскольку они не играют определяющей роли в построении системы качества лаборатории в целом.

*Градуировка в методиках измерений и расчет их результатов (2, 3)* во многих случаях компьютеризируются вместе с аналитическим прибором. Однако в большинстве лабораторий применяются (и будут применяться в обозримом будущем) методики, в которых не используются компьютеризированные приборы. В таких случаях соответствующие расчеты выполняются отдельно, и, хотя их трудоемкость не слишком велика, использование вычислительной техники и в этих случаях желательно из-за особенностей обычно используемого при градуировке метода наименьших квадратов Гаусса (см. разд. 2.7.2.2).

*Оценку метрологических характеристик методик, проверку приемлемости и контроль стабильности результатов измерений (4–6)*, см. разд. 4.5) крайне трудно выполнять без использования компьютера. Помимо достаточно сложных статистических расчетов, при которых необходимо искать табличные значения критических значений статистических функций и приписанные характеристики методик, необходимо вести множество контрольных карт, при этом отслеживая состояние каждой из них, и периодически уточнять нанесенные на карты контрольные пределы (см. выше, разд. 4.5.1.3 и сл.). Подвижные попытки делать это вручную (а карты рисовать на миллиметровой бумаге) почти всегда заканчиваются неудачей. При проверке приемлемости результатов измерений по современным алгоритмам (см. выше, разд. 4.3.3) без компьютера необходимо постоянно использовать калькулятор, держа в памяти показатели повторяемости (сходимости) для разных диапазонов определяемых концентраций используемых методик. Также сложно без компьютера использовать и *другие способы внутрилабораторного контроля (7)*, основанные на оценке состояния измерений за какой-то период (простейший оперативный, «статистический» контроль по альтернативному признаку, контроль с применением метода «введено-найдено» и т. д., описанных в разд. 4.5.2).

*Регистрация поступающих в лабораторию проб, их кодирование, отслеживание движения пробы в лаборатории (особенно в случаях, когда для ее исследования приме-*



няются разные методики КХА), подготовка протоколов (8, 9) — не слишком сложные процедуры при их выполнении вручную. Как показывает практика, количество так называемых пре- и постаналитических ошибок (происходящих на этапах поступления проб в лабораторию и выписки результатов), при использовании бумажных форм ведения этой документации невелико. Более того, среди практиков бытует мнение, что дополнительный контроль, который автоматически происходит, например, при выписывании результатов, даже полезен. Однако с ростом числа выполняемых исследований быстро растет как число ошибок (а они совершенно недопустимы), так и трудоемкость этих операций, что делает целесообразной и экономически выгодной компьютеризацию этих процедур. Особенно резкий скачок в производительности труда происходит при переходе на кодирование проб магнитными либо штрих-кодowymi метками и последующем автоматическом считывании этой информации при измерениях, однако это требует больших затрат и если и имеет смысл, то только в очень больших лабораториях с огромным потоком проб.

*Ведение лабораторной документации, связанной с системой качества лаборатории, хранение и печать содержащейся в ней информации (10–18)* кажется не слишком трудоемким. На самом деле те, кто пытался это делать, знают, как это непросто. Более того, сложность этой работы резко возрастает при увеличении размеров лаборатории и расширении перечня выполняемых исследований. Основную сложность представляет необходимость централизованного ведения достаточно большого числа журналов при децентрализации многих функций (работа с документами, учет реактивов и стандартных образцов, поверка приборов и т. д.). Отдельная проблема — необходимость своевременности закупки свежих реактивов, поверке средств измерений, обучению сотрудников и пр. Заметим, что трудоемкость этих видов деятельности резко возрастает при внедрении современных требований [138, 139]. На практике даже в не очень больших лабораториях для выполнения этих функций вручную приходится выделять специальных сотрудников. Использование компьютерных программ, особенно сетевых версий, позволяет резко упростить эту «бюрократическую» деятельность и распределить соответствующие функции между сотрудниками лаборатории оптимальным образом.

Таким образом, практически все виды деятельности современной лаборатории, выполняющей рутинные исследования, в той или иной степени поддаются компьютеризации, и процесс оснащения лабораторий компьютерной техникой идет весьма активно. Рассмотрим состояние и основные тенденции этого процесса.

#### 4.7.2. Основные пути компьютеризации лабораторий.

Прежде всего отметим, что доля лабораторий, использующих компьютерную технику тем или иным способом, в любом виде, близка к 80% и продолжает расти. Для дальнейшего анализа была использована классификация компьютерных программ, показанная на рис. 4.9. Хотя такая классификация, как и любая другая, является условной, она весьма удобна для рассмотрения проблемы.

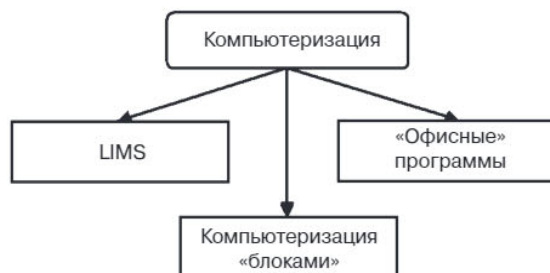


Рис. 4.9. Классификация программ для компьютеризации лабораторий

**LIMS (Laboratory Information Management Systems)**<sup>103</sup> — специализированные программные (иногда — программно-аппаратные) комплексы, при разработке которых ставится задача охватить все стороны деятельности лабораторий. При этом часто такие системы содержат не только собственно лабораторную часть, но и интегрируются в более крупные АСУ (например, заводские), либо содержат функции передачи информации в другие подразделения (например, технологом на производстве). Иногда LIMS предполагают непосредственное подключение к ним аналитических приборов.

**Компьютеризация «блоками»**<sup>104</sup> предполагает наличие в лаборатории *специализированных* профессиональных программ. Это могут быть программы для выполнения расчетов, ведения документации и т. д. Основное отличие этой группы программ от LIMS — отсутствие универсальности. Каждая из них предназначена для решения определенного круга задач, возникающих в лабораториях.

**«Офисные» программы** названы нами так потому, что они либо входят в состав пакета Microsoft office (Microsoft Corporation, США) — наиболее распространенного в России — либо созданы средствами этих программ самими лабораториями. Так, программу Microsoft Word часто используют при подготовке протоколов исследований (предварительно создав соответствующие шаблоны), на основе приложений к Microsoft Excel создают мини-программу для ведения контрольных карт и т. д. Сюда же можно отнести относительно простые самодельные программы, изготавливаемые в конкретных лабораториях для внутренних целей.

#### 4.7.3. Достоинства и недостатки разных классов программ

**LIMS.** Литература по этому виду лабораторных компьютерных программ весьма обширна (см., например, [141–146]). На отечественном рынке представлены как адаптированные зарубежные «LabWare LIMS» фирмы LabWare, «SampleManager LIMS» фирмы Thermo Fisher Scientific и другие, так и отечественные программы (например, ЛИС «Химик-аналитик, НИИ ВН, Томск). Все они претендуют на охват всех или большинства вышеперечисленных видов лабораторной деятельности.

<sup>103</sup> В отечественной литературе используются аббревиатуры ЛИС (лабораторные информационные системы) и, реже, ЛИУС (лабораторные информационно-управляющие системы).

<sup>104</sup> Иногда специализированные программы этой группы называют также «mini-LIMS», программами типа «Лабораторный журнал» и т. д.



Программы имеют множество блоков, при взаимодействии которых возникает огромное число связей между ними. Все это не только приводит к очень сложным программам, но даже заставляет использовать специальную терминологию, не используемую обычно в лабораториях.

Вследствие этого LIMS, преимуществом которых является полная или почти полная компьютеризация деятельности лаборатории, имеют существенные недостатки. Это:

- высокая начальная стоимость системы (реально начинающаяся от \$80 000 и доходящая до нескольких сот тысяч долларов, данные 2014 г.);
- обязательность обучения сотрудников лаборатории пользованию программой, поскольку обычного владения компьютером на уровне пользователя для работы с ней недостаточно;
- необходимость существенной перенастройки программы при изменениях в работе лаборатории, например, при появлении новых методов исследований. Почти всегда при этом необходимо участие специалистов фирмы-разработчика, что выливается в существенные дополнительные расходы.

По мнению разработчиков LIMS, высокие затраты на их установку и эксплуатацию могут окупаться за счет следующих факторов [144]:

- сокращение персонала за счет более эффективного использования оставшихся сотрудников («перебрасывания» временно свободных работников на «узкие» места, которые можно оперативно выявлять с помощью LIMS);
- оптимизация использования оборудования.

Кроме того, часто утверждается, что при использовании LIMS уменьшается количество ошибок при вводе данных.

Однако практика показывает, что ожидаемой экономии после внедрения LIMS не происходит. В стабильно работающей лаборатории не происходит сокращения персонала, так как все и так более или менее загружены (а имеющаяся специализация исполнителей не позволяет абсолютно свободно перебрасывать их с одного вида исследований на другой). Экстренные ситуации успешно разрешаются опытными руководителями среднего звена без всяких компьютеров. Не удастся и уменьшить количество приборов — измерения все равно надо проводить, есть LIMS или нет. Даже если в результате внедрения LIMS в очень крупных лабораториях и удастся сократить рядового сотрудника, это практически не повлияет на стоимость исследований; такая ситуация сохранится и в обозримом будущем. В то же время эксплуатация LIMS обычно требует специального персонала.

Отдельно необходимо рассмотреть вопрос об уменьшении ошибок ввода. Действительно, при непосредственном подключении всех приборов к LIMS (что чрезвычайно дорого и не всегда возможно в принципе) вводить данные вручную не надо. Однако количество ошибок при вводе только первичных данных, записанных на бумаге, невелико, и они в большинстве случаев выявляются при проверке приемлемости результатов измерений либо при подготовке протоколов (этот этап, по сути, является дополнительной формой контроля — опытный сотрудник может обнаружить грубые ошибки, не замеченные ранее). Кроме того, во многих случаях (например, в фармакологических лабораториях) фиксация первичных результатов измерений на бумаге с подписью исполнителя обязательна.



Из вышесказанного очевидно, что внедрение LIMS в абсолютном большинстве случаев достаточно сложно и экономически не эффективно. Вероятно, это и является причиной того, что их реально использует небольшая часть лабораторий. Причем во всех случаях это — лаборатории крупных предприятий, то есть имеющих богатого «спонсора».

Специализированные профессиональные программы (компьютеризация «блоками»). Такие программы весьма распространены. Как уже было сказано, каждая из них предназначена для решения определенной части перечисленных выше видов деятельности лабораторий.

Разные виды деятельности в лаборатории можно объединить в группы («блоки») с учетом их специфики. Задачи, решаемы в каждой группе (блоке) задач, обычно выполняются одними и теми же людьми. Поэтому они, естественно, «естественно» компьютеризируются в рамках одной программы (рис. 4.10). При этом, конечно, учитывается в первую очередь удобство использования каждой программой, а не простота ее написания. Стоимость отечественных программ этого класса лежит в диапазоне \$800–2000 (данные 2019 г.).

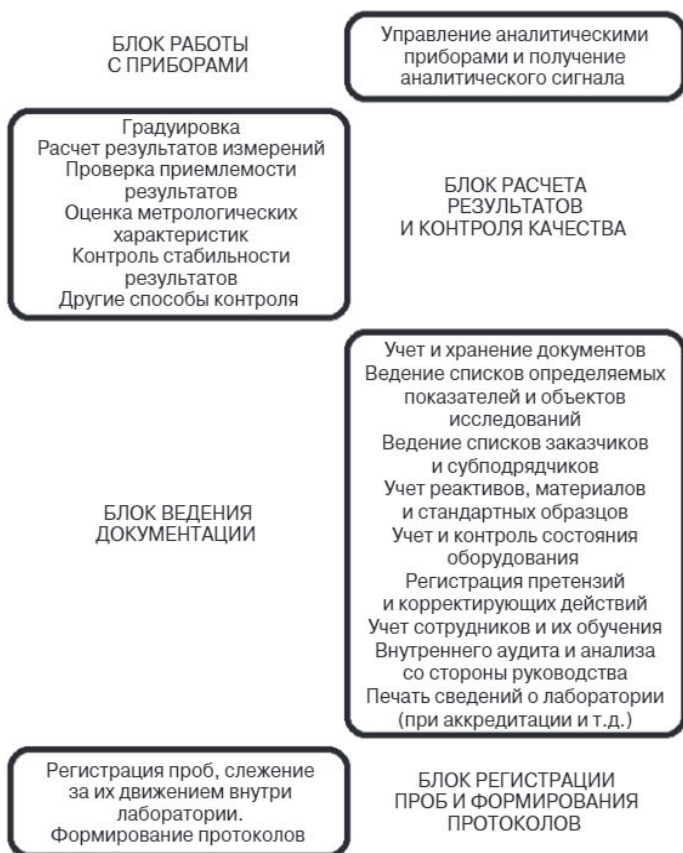


Рис. 4.10. Компьютеризация «блоками». Задачи каждого блока, «естественно» решаемые в рамках одной программы, включены в один прямоугольник

К сожалению, литература по специализированным программам весьма разрозненна и не информативна. Рассмотрим возможности этого пути для каждого блока задач в общем виде и на примере наиболее распространенных в России программ DControl и QControl с блоком SControl). При этом мы не будем обсуждать *работу с приборами*, поскольку соответствующие программы являются составной частью этих приборов (см. выше). Заметим только, что часто такие программы приобретаются отдельно у независимых производителей, например, при компьютеризации хроматографов (например, [147, 148]).

*Ведение лабораторной документации* полностью охватывается программой DControl (Document Control). Поскольку перечисленные в описании этого блока задачи в основном соответствуют ГОСТ ИСО/МЭК 17025 [73], программа имеет несколько разделов, комплиментарных разделам этого документа (рис. 4.11). С помощью программы DControl удастся решить все основные задачи, связанные с документацией (см. выше), включая создание форм печати, содержащих сведения по всем разделам программы и соответствующих требованиям внешних организаций (органы по аккредитации, вышестоящие и контролирующие организации). Другими словами, на основе программы DControl можно построить систему качества лаборатории. Примеров таких систем достаточно много (программа DControl используется в нескольких сотнях лабораторий).

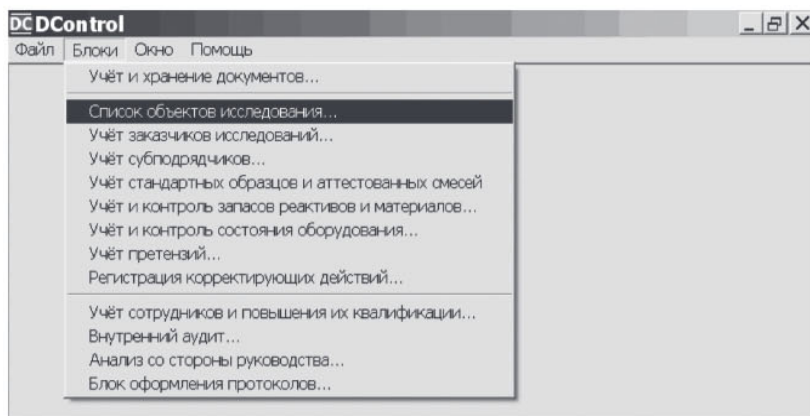


Рис. 4.11. Некоторые блоки программы DControl

Надо особо отметить, что использование программы DControl в сетевом варианте позволяет распределить обязанности по ведению документации, разным видам учета, внутренним проверкам, ведению отчетности и т. д. среди всех сотрудников лаборатории, что позволяет отказаться от выделения специалистов, ответственных за качество и освобожденных от другой работы, и разгрузить руководство лаборатории. Нам неизвестны другие программы, специально предназначенные для решения этого блока задач.

*Задачи расчета результатов и контроля качества* решаются с помощью программы QControl [149–151]. Основной ее частью является электронный лабораторный журнал (блок-схема которого показана на рис. 4.12), в который вводятся первичные

данные (аналитический сигнал при химическом анализе). Предварительно в программу заносятся сведения о методиках: число измерений для рутинных проб, способ расчета окончательно приводимого результата, стандартные отклонения повторяемости (сходимости) и воспроизводимости, способы проверки приемлемости результатов и контроля стабильности и т. д. Программа автоматически рассчитывает окончательно приводимые результаты измерений по градуировочной зависимости или (и) по формуле, проводит проверку приемлемости каждого результата и контроль стабильности измерений. Кроме того, с помощью программы QControl можно проводить и другие расчеты, которые могут оказаться необходимыми в лаборатории.



Рис. 4.12. Блок-схема работы Лабораторного журнала программы QControl

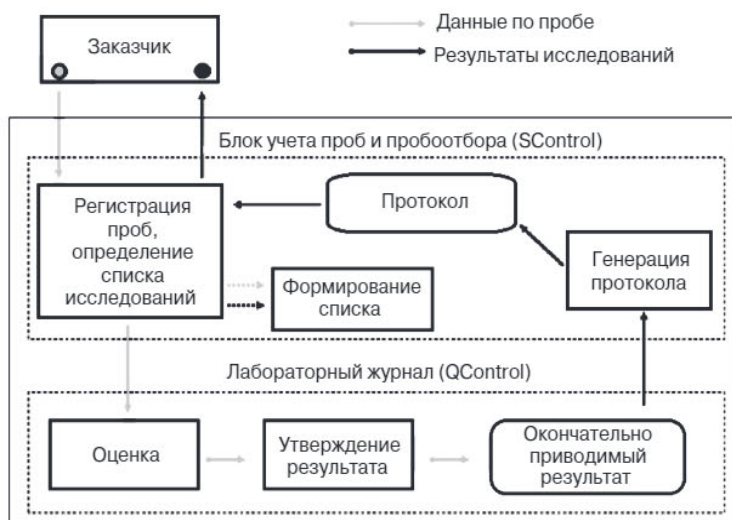


Рис. 4.13. Взаимодействие блока SControl и программы QControl



*Регистрация проб, слежение за их движением внутри лаборатории и формирование протоколов* выполняются с помощью блока SControl. Фактически это отдельная программа, взаимодействующая с программой QControl через специальный программный интерфейс (см. рис. 4.13). Такой подход обусловлен как различием выполняемых этими программами задач, так и тем, что эти задачи решаются разными людьми. Подробнее об описанных программах см. [152].

### **«Офисные» программы**

Как указано выше, сюда можно отнести два основных вида программ:

- стандартные программы, входящие в пакет Microsoft Office (Word, Excel и др.);
- самодельные программы, созданные в лабораториях. Если раньше такие программы писались с помощью базовых инструментов программирования (языки программирования общего назначения C++, Pascal и т. д.), то сейчас чаще используется приложение к Microsoft Excel.

Нет смысла обсуждать здесь использование в лабораториях собственно офисных программ — они общеизвестны, а их использование для подготовки разнообразных документов (Word) и целей учета (Excel) очевидно. Следует отметить, однако, что эти программы гораздо менее удобны, чем специализированные программы, используемые при компьютеризации «блоками» — без серьезной модификации они охватывают лишь часть задач, возникающих в лаборатории.

Самодельные программы, написанные в лабораториях, чаще всего решают весьма узкие задачи (например, ведение контроля стабильности с помощью контрольных карт). Обычно разработку таких программ оправдывают их дешевизной. Однако это самообман — работа программиста в любом случае стоит немалых денег, и на поверку оказывается, что стоимость таких «бесплатных» программ намного превышает стоимость специализированных программ, разработанных для компьютеризации «блоками». Но не это главное. Для того чтобы сделать удобную и надежную программу, лабораторный специалист должен пройти весь путь, который проходят разработчики коммерческих программ. Этого никогда не происходит, и кустарные «офисные» программы по качеству сильно уступают специализированным коммерческим программам. Действительно, никому не приходит в голову, например, разрабатывать для собственных нужд оригинальный фотометр — несмотря на то, что в продаже есть все нужные компоненты, а в лаборатории инженер.

Следствием этого является процесс «вымывания» самодельных программ из обихода лабораторий — пик использования их пришелся на 1990-е годы, и с тех пор число таких программ быстро уменьшается.

Таким образом, проведенный анализ рассмотренных классов программ, используемых при построении системы обеспечения качества аналитических лабораторий, позволяет прийти к следующим основным выводам.

«Офисные» и простые кустарные программы постепенно вымываются из обихода. Конечно, использование основных программ пакета Microsoft Office (Word, Excel и др.) в обозримом будущем не прекратится, но будет носить ограниченный характер.

LIMS в обозримом будущем, как и сейчас, будут иметь весьма ограниченное применение — в лабораториях, имеющих богатого «спонсора» (главным образом лаборатории крупных успешных предприятий), в тех случаях, когда требование полной компьютеризации лаборатории в рамках одной программы (взаимосвязанного комплекса программ) вызвано производственной необходимостью.

Основным путем компьютеризации аналитических и других лабораторий останется использование специализированных программ (компьютеризация «блоками»), позволяющих полностью удовлетворять современным требованиям к системе обеспечения качества лабораторий при минимальных затратах финансовых и трудовых ресурсов.

# Приложение 1. Единицы системы СИ

## 1. Основные единицы системы СИ

Величина		Единица			
Наименование	Размер-ность	Наименование	Обозначение		Определение
			Русское	Между-народное	
1	2	3	4	5	6
Длина	L	метр	м	m	Метр есть длина пути, проходящего светом в вакууме за 1/299 792 458 с
Масса	M	килограмм	кг	kg	Килограмм есть постоянная Планка, поделенная на 6,626070040×10 <sup>-34</sup> м <sup>2</sup> ·с <sup>-1</sup>
Время	T	секунда	с	s	Секунда есть время, равное 9 192 631 770 периодов излучения, соответствующего переходу между двумя сверхтонкими уровнями основного состояния атома цезия-133 при 0 К
Сила электрического тока	I	Ампер	A	A	Ампер есть электрический ток, соответствующий потоку 1/1,6021766208×10 <sup>-19</sup> элементарных электрических зарядов в секунду
Термодинамическая температура	Θ	Кельвин	К	K	Кельвин есть единица термодинамической температуры. Она соответствует изменению тепловой энергии на 1,38064852×10 <sup>-23</sup> джоулей
Количество вещества	N	Моль	моль	mol	Моль есть количество вещества системы, которая содержит 6,022140857×10 <sup>23</sup> специфицированных структурных единиц. <i>Примечание:</i> структурные элементы должны быть специфицированы и могут быть атомами, молекулами, ионами, электронами и другими частицами или специальными группами частиц
Сила света	J	Кандела	кд	cd	Кандела есть сила света в заданном направлении источника, испускающего монохроматическое излучение частотой 540·10 <sup>12</sup> Гц, энергетическая сила излучения которого в этом направлении равна 1/683 Вт/ср



## 2. Производные единицы системы СИ широкого распространения

### 2.1. Пространство и время

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Плоский угол	I	Радян	рад	rad	Радян есть угол между двумя радиусами окружности, длина дуги между которыми равна одному радиусу
Телесный угол	I	Стерadian	ср	sr	Стерadian есть телесный угол с вершиной в центре сферы, вырезающий на поверхности сферы площадь, равную площади квадрата, длина сторон которого равна радиусу сферы
Площадь	L <sup>2</sup>	квадратный метр	м <sup>2</sup>	m <sup>2</sup>	Квадратный метр есть площадь квадрата со сторонами, равными 1 м
Объем	L <sup>3</sup>	кубический метр	м <sup>3</sup>	m <sup>3</sup>	Кубический метр есть объем куба с ребрами, длина которых равна 1 м
Линейная скорость	L T <sup>-1</sup>	метр в секунду	м/с	m/s	Метр в секунду есть скорость прямолинейно и равномерно движущейся точки, при которой эта точка за время 1 с перемещается на расстояние 1 м
Линейное ускорение	L T <sup>-2</sup>	метр на секунду в квадрате	м/с <sup>2</sup>	m/s <sup>2</sup>	Метр на секунду в квадрате есть ускорение прямолинейно и равноускоренно движущейся точки, при котором за время 1 с скорость ее изменяется на 1 м/с
Угловая скорость	T <sup>-1</sup>	радиан в секунду	рад/с	rad/s	Радян в секунду есть угловая скорость равномерно вращающейся точки, при которой за время 1 с точка совершает поворот относительно оси вращения на угол 1 рад
Угловое ускорение	T <sup>-2</sup>	радиан на секунду в квадрате	рад/с <sup>2</sup>	rad/s <sup>2</sup>	Радян на секунду в квадрате есть угловое ускорение равноускоренно вращающейся точки, при которой за время 1 с угловая скорость ее увеличивается на 1 рад/с

2.2. Периодические и связанные с ними явления

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размер- ность	Наименование	Обозначение		Определение
			Русское	Между- народное	
1	2	3	4	5	6
Частота	$T^{-1}$	Герц	Гц	Hz	Герц есть частота периодического процесса, при которой за время, равное 1 с, совершается один цикл периодического процесса
Частота вращения	$T^{-1}$	секунда в минус первой степени	$c^{-1}$	$s^{-1}$	Секунда в минус первой степени есть частота равномерного вращения, при которой за время, равное 1 с, тело совершает один полный оборот
Частота дискретных событий	$T^{-1}$	секунда в минус первой степени	$c^{-1}$	$s^{-1}$	Секунда в минус первой степени есть частота дискретных событий, при которой за время, равное 1 с, тело происходит одно дискретное событие
Круговая (циклическая) частота	$T^{-1}$	секунда в минус первой степени	$c^{-1}$	$s^{-1}$	Секунда в минус первой степени есть частота круговая частота, при которой за время, равное 1 с, совершается $2\pi$ циклов вращения
Длина волны	L	метр	м	m	Метр есть длина волны, при которой частота колебаний 1 Гц соответствует скорости распространения волны 1 м/с
Волновое число	$L^{-1}$	метр в минус первой степени	$m^{-1}$	$m^{-1}$	Метр в минус первой степени есть волновое число колебаний с длиной волны 1 м

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Плотность (массы)	$L^{-3}M$	килограмм на кубический метр	кг/м <sup>3</sup>	kg/m <sup>3</sup>	Килограмм на кубический метр есть плотность однородного вещества, масса которого при объеме 1 м <sup>3</sup> равна 1 кг
Линейная плотность	$L^{-1}M$	килограмм на метр	кг/м	kg/m	Килограмм на метр есть линейная плотность тела, масса которого при длине равной 1 м равна 1 кг
Поверхностная плотность	$L^{-2}M$	килограмм на квадратный метр	кг/м <sup>2</sup>	kg/m <sup>2</sup>	Килограмм на квадратный метр есть поверхностная плотность тела, масса которого при площади 1 м <sup>2</sup> равна 1 кг
Удельный объем	$L^3M^{-1}$	кубический метр на килограмм	м <sup>3</sup> /кг	m <sup>3</sup> /kg	Кубический метр на килограмм есть удельный объем однородного вещества, объем которого при массе 1 кг равен 1 м <sup>3</sup>
Импульс (количество движения)	$LM T^{-1}$	килограмм-метр в секунду	кг·м/с	kg·m/s	Килограмм-метр в секунду равен импульсу материальной точки массой 1 кг, движущейся со скоростью 1 м/с
Момент инерции (динамический)	$L^2M$	килограмм-метр в квадрате	кг·м <sup>2</sup>	kg·m <sup>2</sup>	Килограмм-метр в квадрате есть момент инерции материальной точки с массой 1 кг, находящейся на расстоянии 1 м от оси вращения
Сила. Вес	$LM T^{-2}$	ньютон	Н	N	Ньютон есть сила, сообщаящая телу массой 1 кг ускорение 1 м/с <sup>2</sup> в направлении действия силы. Он же — единица веса и силы тяжести СИ
Момент силы	$L^2MT^{-2}$	ньютон-метр	Н·м	N·m	Ньютон-метр есть момент силы, создаваемый силой 1 Н относительно точки, расположенной на расстоянии 1 м от оси вращения
Импульс силы	$LM T^{-1}$	ньютон-секунда	Н·с	N·s	Ньютон-секунда есть импульс силы, вызываемый силой, равной 1 Н, действующей в течение 1 с
Давление	$L^{-1}MT^{-2}$	паскаль	Па	Pa	Паскаль есть давление, вызываемое силой 1 Н, равномерно распределенной по поверхности площадью 1 м <sup>2</sup> , расположенной перпендикулярно направлению действия силы
Нормальное механическое напряжение	$L^{-1}MT^{-2}$	паскаль	Па	Pa	Паскаль есть нормальное механическое напряжение, вызываемое упругой силой 1 Н при равномерном распределении ее по сечению площадью 1 м <sup>2</sup> , расположенному перпендикулярно направлению действия силы



2.3. Механика (окончание)

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размерность	Наименование	Обозначение		Определение
1	2	3	Русское	Международное	6
Модуль продольной упругости (Юнга)	$L^{-1}MT^{-2}$	паскаль	Па	Pa	Паскаль есть модуль продольной упругости тела, испытывающего относительное удлинение, равное 1, при нормальном механическом напряжении 1 Па
Вязкость (динамическая)	$L^{-1}MT^{-1}$	паскаль-секунда	Па·с	Pa·s	Паскаль-секунда есть динамическая вязкость среды, тангенциальное напряжение в которой при ламинарном движении и при разности скоростей слоев, находящихся на расстоянии 1 м по нормали к направлению скорости, 1 м/с, равно 1 Па
Вязкость кинематическая	$L^2T^{-1}$	квадратный метр в секунду	м <sup>2</sup> /с	m <sup>2</sup> /s	Квадратный метр в секунду есть кинематическая вязкость, при которой динамическая вязкость среды плотностью 1 кг/м <sup>3</sup> равна 1 Па·с
Поверхностное натяжение	$MT^{-2}$	ньютон на метр	Н/м	N/m	Ньютон на метр есть поверхностное натяжение, создаваемое силой 1 Н, прикладываемой к участку контура свободной поверхности длиной 1 м и действующей нормально к контуру и по касательной к поверхности
Работа. Энергия. Потенциальная энергия. Кинетическая энергия	$L^2MT^{-2}$	джоуль	Дж	J	Джоуль есть работа, совершаемая силой 1 Н при перемещении точки приложения силы на 1 м в направлении действия силы. Джоуль является также единицей энергии системы СИ
Мощность	$L^2MT^{-3}$	ватт	Вт	W	Ватт есть мощность, при которой за время 1 с совершается работа 1 Дж
Объемный расход	$L^3T^{-1}$	кубический метр в секунду	м <sup>3</sup> /с	m <sup>3</sup> /s	Кубический метр в секунду есть объемный расход, при котором сквозь определенное сечение за время 1 с равномерно перемещается вещество объемом 1 м <sup>3</sup>
Массовый расход	$MT^{-1}$	килограмм в секунду	кг/с	kg/s	Килограмм в секунду есть массовый расход, при котором сквозь определенное сечение за время 1 с равномерно перемещается вещество массой 1 кг

2.4. Теплота

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Температурный коэффициент линейного расширения	$\Theta^{-1}$	кельвин в минус первой степени	$K^{-1}$	$K^{-1}$	Кельвин в минус первой степени есть температурный коэффициент линейного расширения, при котором изменение температуры на 1 К (по отношению к начальной) вызывает относительное изменение длины, равное 1
Температурный коэффициент объемного расширения	$\Theta^{-1}$	кельвин в минус первой степени	$K^{-1}$	$K^{-1}$	Кельвин в минус первой степени есть температурный коэффициент объемного расширения, при котором изменение температуры на 1 К (по отношению к начальной) вызывает относительное изменение объема, равное 1
Температурный коэффициент давления	$\Theta^{-1}$	кельвин в минус первой степени	$K^{-1}$	$K^{-1}$	Кельвин в минус первой степени есть температурный коэффициент давления, при котором изменение температуры на 1 К (по отношению к начальной) вызывает относительное изменение давления, равное 1
Температурный градиент	$L^{-1}\Theta$	кельвин на метр	$K/m$	$K/m$	Кельвин на метр есть температурный градиент поля, в котором на участке длиной 1 м температура в направлении градиента температур меняется на 1 К
Количество теплоты	$L^2MT^{-2}$	джоуль	Дж	J	Джоуль есть количество теплоты, эквивалентное работе 1 Дж
Удельное количество теплоты	$L^2T^{-2}$	джоуль на килограмм	Дж/кг	J/kg	Джоуль на килограмм есть удельное количество теплоты процесса, при котором к веществу с массой 1 кг подводится (отводится) количество теплоты 1 Дж
Теплоемкость	$L^2MT^{-2}\Theta^{-1}$	джоуль на кельвин	Дж/К	J/K	Джоуль на кельвин есть теплоемкость тела, повышающего температуру на 1 К при сообщении ему количества теплоты 1 Дж
Удельная теплоемкость	$L^2T^{-2}\Theta^{-1}$	джоуль на килограмм-кельвин	Дж/(кг·К)	J/(kg·K)	Джоуль на килограмм-кельвин есть удельная теплоемкость вещества, имеющего при массе равной 1 кг теплоемкость 1 Дж/К
Тепловой поток	$L^2MT^{-3}$	Ватт	Вт	W	Ватт есть тепловой поток, при котором за 1 с подводится (или отводится) количество теплоты 1 Дж
Поверхностная плотность теплового потока	$MT^{-3}$	ватт на квадратный метр	Вт/м <sup>2</sup>	W/m <sup>2</sup>	Ватт на квадратный метр есть поверхностная плотность теплового потока 1 Вт, равномерно распределенного на поверхности с площадью 1 м <sup>2</sup>
Теплопроводность	$LM^{-1}T^{-3}\Theta^{-1}$	ватт на метр-кельвин	Вт/(м·К)	W/(m·K)	Ватт на метр-кельвин есть теплопроводность вещества, в котором при стационарном режиме с поверхностной плотностью теплового потока 1 Вт/м устанавливается температурный градиент 1 К/м

2.5. Электричество и магнетизм

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Электрический заряд (количество электричества)	Тл	Кулон	Кл	С	Кулон есть количество электричества, проходящее сквозь поперечное сечение проводника при силе тока 1 А за время 1 с
Плотность (пространственная) электрических зарядов	$L^{-3}Тл$	кулон на метр кубический	Кл/м <sup>3</sup>	С/м <sup>3</sup>	Кулон на метр кубический есть пространственная плотность электрического заряда, при которой заряд, равномерно распределенный в пространстве с объемом 1 м <sup>3</sup> , равен 1 Кл
Плотность (поверхностная) электрических зарядов	$L^{-2}Тл$	кулон на квадратный метр	Кл/м <sup>2</sup>	С/м <sup>2</sup>	Кулон на квадратный метр есть поверхностная плотность электрического заряда, при которой заряд, равномерно распределенный по поверхности площадью 1 м <sup>2</sup> , равен 1 Кл
Линейная плотность электрических зарядов	$L^{-1}Тл$	кулон на метр	Кл/м	С/м	Кулон на метр есть линейная плотность электрического заряда, при которой заряд, равномерно распределенный по линии длиной 1 м, равен 1 Кл
Поверхностная плотность электрического тока	$L^{-2}л$	ампер на квадратный метр	А/м <sup>2</sup>	А/м <sup>2</sup>	Ампер на квадратный метр есть поверхностная плотность электрического тока, при которой сила тока, равномерно распределенного по поперечному сечению проводника площадью 1 м <sup>2</sup> , равна 1 А
Линейная плотность электрического тока	$L^{-1}л$	ампер на метр	А/м	А/м	Ампер на метр есть линейная плотность электрического тока, при которой сила тока, равномерно распределенного по сечению тонкого листового проводника шириной 1 м, равна 1 А
Электрический потенциал. Разность потенциалов. Электрическое напряжение	$L^2М.Т^{-3}л^{-1}$	вольт	В	В	Вольт есть электрический потенциал в данной точке электрического поля, при котором работа по перемещению положительного заряда 1 Кл из данной точки в точку с нулевым потенциалом равна 1 Дж. Вольт есть электрическое напряжение между двумя точками внутри проводника, при котором работа по перемещению положительного заряда 1 Кл между этими двумя точками равна 1 Дж



## 2.5. Электричество и магнетизм (окончание)

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размер- ность	Наименование	Обозначение		Определение
			Русское	Между- народное	
1	2	3	4	5	6
Электродвижу- щая сила	$L^2 M \cdot T^{-3} I^{-1}$	вольт	В	V	Вольт есть электродвижущая сила источника тока, при которой совершается работа сторонними силами по перемещению положительного заряда 1 Кл от отрицательного к положительному полюсу источника вдоль всей электрической цепи равная 1 Дж
Напряженность электрического поля	$L M \cdot T^{-3} I^{-1}$	вольт на метр	В/м	V/m	Вольт на метр есть напряженность электрического поля, при которой поле действует на точечный заряд 1 Кл силой 1 Н
Электрическая емкость	$L^{-2} M^{-1} \cdot T^4 I^2$	фарад	Ф	F	Фарад есть электрическая емкость конденсатора, при которой заряд 1 Кл создает между обкладками его напряжение 1 В
Электрическое сопротивление	$L^2 M \cdot T^{-3} I^{-2}$	ом	Ом	W	Ом есть электрическое сопротивление проводника, между концами которого возникает напряжение 1 В при силе постоянного тока 1 А
Удельное электрическое сопротивление	$L^3 M \cdot T^{-3} I^{-2}$	ом-метр	Ом·м	$\Omega \cdot m$	Ом-метр есть удельное электрическое сопротивление материала с площадью попереч- ного сечения 1 м <sup>2</sup> и длиной 1 м, имеющего сопротивление 1 Ом
Электрическая проводимость	$L^{-2} M^{-1} \cdot T^3 I^2$	сименс	См	S	Сименс есть электрическая проводимость участка электрической цепи сопротивлением 1 Ом
Удельная электрическая проводимость	$L^{-3} M^{-1} \cdot T^3 I^2$	сименс на метр	см/м	S/m	Сименс на метр есть удельная электрическая проводимость материала проводника, при которой удельное электрическое сопротивление проводника равно 1 Ом·м
Магнитный поток (поток магнитной индукции)	$L^2 M \cdot T^{-2} I^{-1}$	вебер	Вб	Wb	Вебер есть магнитный поток, при убывании которого до нуля в сцепленной с ним элект- рической цепи сопротивлением 1 Ом сквозь поперечное сечение проводника проходит количество электричества 1 Кл

## 2.6. Акустика

ВЕЛИЧИНА		ЕДИНИЦА		
Наименование	Размерность	Наименование	Обозначение	
			Русское	Международное
1	2	3	4	5
Плотность звуковой энергии	$L^{-1}MT^{-2}$	джоуль на метр кубический	Дж/м <sup>3</sup>	J/m <sup>3</sup>
Поток звуковой энергии (звуковая мощность)	$L^2MT^{-3}$	ватт	Вт	W
Интенсивность звука	$MT^{-3}$	ватт на метр квадратный	Вт/м <sup>2</sup>	W/m <sup>2</sup>
		6		
		Джоуль на метр кубический есть плотность звуковой энергии в канале объемом 1 м <sup>3</sup> при звуковой энергии 1 Дж		
		Ватт есть поток звуковой энергии, при котором за время 1 с сквозь площадь, перпендикулярную направлению распространения волн, переносится звуковая энергия 1 Дж		
		Ватт на метр квадратный есть интенсивность звука при потоке звуковой энергии 1 Вт и площади поперечного сечения канала 1 м <sup>2</sup>		

## 2.7. Оптическое излучение

ВЕЛИЧИНА		ЕДИНИЦА		
Наименование	Размерность	Наименование	Обозначение	
			Русское	Международное
1	2	3	4	5
Плотность (объемная) энергии излучения	$L^{-1}MT^{-2}$	джоуль на метр кубический	Дж/м <sup>3</sup>	J/m <sup>3</sup>
Лучистая экспозиция	$MT^{-2}$	джоуль на квадратный метр	Дж/м <sup>2</sup>	J/m <sup>2</sup>
Поток излучения (мощность излучения)	$L^2MT^{-3}$	ватт	Вт	W
Плотность (поверхностная) потока излучения (лучистого потока)	$MT^{-3}$	ватт на метр квадратный	Вт/м <sup>2</sup>	W/m <sup>2</sup>
		6		
		Джоуль на метр кубический есть плотность лучистой энергии, при которой тело с объемом 1 м <sup>3</sup> излучает энергию 1 Дж		
		Джоуль на квадратный метр есть лучистая экспозиция, при которой на поверхность 1 м <sup>2</sup> падает излучение 1 Дж		
		Ватт есть поток излучения, при котором за время 1 с излучается энергия 1 Дж		
		Ватт на метр квадратный есть поверхностная плотность потока излучения, при которой поверхность площадью 1 м <sup>2</sup> излучает (или поглощает) поток излучения 1 Вт		

## 2.7. Оптическое излучение (окончание)

ВЕЛИЧИНА			ЕДИНИЦА		
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Облученность	$\text{MT}^{-3}$	ватт на метр квадратный	$\text{Вт}/\text{м}^2$	$\text{W}/\text{m}^2$	Ватт на метр квадратный есть облученность, при которой на поверхность площадью $1 \text{ м}^2$ падает поток излучения $1 \text{ Вт}$
Сила излучения	$\text{L}^2\text{MT}^{-3}$	ватт на стерадиан	$\text{Вт}/\text{ср}$	$\text{W}/\text{ср}$	Ватт на стерадиан есть сила излучения точечного источника, излучающего в телесном угле $1 \text{ ср}$ поток излучения $1 \text{ Вт}$
Лучистость	$\text{MT}^{-3}$	ватт на стерадиан-квадратный метр	$\text{Вт}/\text{ср}\cdot\text{м}^2$	$\text{W}/\text{sr}\cdot\text{m}^2$	Ватт на стерадиан-квадратный метр есть лучистость равномерно излучающей плоской поверхности площадью $1 \text{ м}^2$ в перпендикулярном к ней направлении, при силе излучения $1 \text{ Вт}/\text{ср}$
Световой поток	J	люмен	лм	lm	Люмен есть световой поток, испускаемый точечным источником силой света $1 \text{ Кд}$ в телесном угле $1 \text{ ср}$
Световая энергия	TJ	люмен-секунда	лм·с	lm·s	Люмен-секунда есть световая энергия светового потока в $1 \text{ лм}$ , действующего в течение времени $1 \text{ с}$
Яркость	$\text{L}^{-3}\text{J}$	кандела на квадратный метр	$\text{Кд}/\text{м}^2$	$\text{cd}/\text{m}^2$	Кандела на квадратный метр есть яркость светящейся поверхности площадью $1 \text{ м}^2$ , при силе света $1 \text{ Кд}$
Светимость	$\text{L}^{-2}\text{J}$	люмен на квадратный метр	$\text{лм}/\text{м}^2$	$\text{lm}/\text{m}^2$	Люмен на квадратный метр есть светимость поверхности площадью $1 \text{ м}^2$ , испускающей световой поток $1 \text{ лм}$
Освещенность	$\text{L}^{-2}\text{J}$	люкс	лк	lx	Люкс есть освещенность поверхности площадью $1 \text{ м}^2$ при падающем на нее световом потоке $1 \text{ лм}$
Световая экспозиция	$\text{L}^{-2}\text{TJ}$	люкс-секунда	лк·с	lx·s	Люкс-секунда есть световая экспозиция, создаваемая за время $1 \text{ с}$ при освещенности $1 \text{ лк}$
Световая эффективность потока излучения	$\text{L}^{-2}\text{M}^{-1}\cdot\text{T}^3\text{J}$	люмен на ватт	лм/Вт	lm/W	Люмен на ватт есть световая эффективность, при которой лучистому потоку $1 \text{ Вт}$ соответствует световой поток $1 \text{ лм}$





2.8. Химия и молекулярная физика

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Молярная масса вещества	$MN^{-1}$	килограмм на моль	кг/моль	kg/mol	Килограмм на моль есть молярная масса вещества, имеющего при количестве вещества 1 моль массу 1 кг
Молярный объем вещества	$L^3N^{-1}$	метр кубический на моль	м <sup>3</sup> /моль	m <sup>3</sup> /mol	Метр кубический на моль есть молярный объем вещества, занимающего при количестве вещества 1 моль объем 1 м <sup>3</sup>
Молярная внутренняя энергия	$L^2M \cdot T^{-2}N^{-1}$	джоуль на моль	Дж/моль	J/mol	Джоуль на моль есть молярная внутренняя энергия вещества, имеющего при количестве вещества 1 моль внутреннюю энергию 1 ДЖ
Концентрация (объемная) молекул	$L^{-3}$	метр в минус первой степени	м <sup>-1</sup>	m <sup>-1</sup>	Метр в минус первой степени есть концентрация (объемная) молекул, при которой в объеме 1 м <sup>3</sup> находится одна молекула
Массовая концентрация компонента	$L^{-3}M$	килограмм на метр кубический	кг/м <sup>3</sup>	kg/m <sup>3</sup>	Килограмм на метр кубический есть массовая концентрация компонента, при которой в объеме смеси (раствора, сплава) 1 м <sup>3</sup> содержится компонент с массой 1 кг
Молярная концентрация компонента	$L^{-3}N$	моль на метр кубический	моль/м <sup>3</sup>	mol/m <sup>3</sup>	Моль на метр кубический есть молярная концентрация компонента в смеси (растворе, сплаве), при которой в объеме смеси (раствора, сплава) 1 м <sup>3</sup> содержится количество вещества 1 моль
Молярность раствора компонента	$M^{-1}N$	моль на килограмм	моль/кг	mol/kg	Моль на килограмм есть молярность раствора, при которой на массу растворителя 1 кг приходится количество растворенного вещества 1 моль
Скорость химической реакции	$L^3T^{-1}N$	моль в секунду на метр кубический	моль/с·м <sup>3</sup>	mol/s·m <sup>3</sup>	Моль в секунду на метр кубический есть средняя скорость одномолекулярной химической реакции, при которой за время 1 с молярная концентрация исходного вещества в растворе изменяется на 1 моль/м <sup>3</sup>

## 2.9. Ионизирующие излучения

ВЕЛИЧИНА			ЕДИНИЦА		
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Энергия ионизирующего излучения	$L^2MT^{-2}$	джоуль	Дж	J	Джоуль есть энергия ионизирующего излучения, эквивалентная работе 1 Дж
Поглощенная доза излучения	$L^2T^{-2}$	грей	Гр	Gy	Грей есть поглощенная доза излучения, при которой облучаемому веществу с массой 1 кг передается энергия ионизирующего излучения 1 Дж
Мощность поглощенной дозы излучения	$L^2T^{-3}$	грей в секунду	Гр/с	Gy/s	Грей в секунду есть мощность поглощенной дозы излучения, при которой за время, равное 1 с облучаемым веществом поглощается доза излучения 1 Гр
Экспозиционная доза рентгеновского и $\gamma$ -излучений	$M^{-1}T$	кулон на килограмм	Кл/кг	C/kg	Кулон на килограмм есть экспозиционная доза рентгеновского и $\gamma$ -излучений, при которых сопряженная корпускулярная эмиссия в сухом атмосферном воздухе массой 1 кг производит ионы, несущие электрический заряд каждого знака, равный 1 Кл
Мощность экспозиционной дозы рентгеновского и $\gamma$ -излучений	$M^{-1}$	ампер на килограмм	A/кг	A/kg	Ампер на килограмм есть мощность экспозиционной дозы рентгеновского и $\gamma$ -излучений, при которой за время 1 с сухому атмосферному воздуху передается экспозиционная доза излучений, равная 1 Кл/кг
Эквивалентная доза излучения	$L^2T^{-2}$	зиверт	Зв	Sv	Зиверт есть эквивалентная доза излучения, при которой поглощенная доза равна 1 Гр, и коэффициент качества излучения равен 1
Мощность эквивалентной дозы излучения	$L^2T^{-3}$	зиверт в секунду	Зв/с	Sv/s	Зиверт в секунду есть мощность эквивалентной дозы излучения, при которой за время 1 с облучаемым веществом поглощается эквивалентная доза излучения 1 Зв
Поток энергии (мощность) ионизирующего излучения	$L^2MT^{-3}$	ватт	Вт	W	Ватт есть поток энергии ионизирующего излучения, при котором за время 1 с сквозь некоторое сечение переносится энергия ионизирующего излучения, равная 1 Дж

2.9. Ионизирующие излучения (окончание)

ВЕЛИЧИНА		ЕДИНИЦА			
Наименование	Размерность	Наименование	Обозначение		Определение
			Русское	Международное	
1	2	3	4	5	6
Интенсивность (плотность потока энергии) ионизирующего излучения	$\text{МТ}^{-3}$	ватт на квадратный метр	$\text{Вт}/\text{м}^2$	$\text{W}/\text{m}^2$	Ватт на квадратный метр есть интенсивность излучения, при которой на поверхность площадью $1 \text{ м}^2$ падает излучение мощностью $1 \text{ Вт}$ перпендикулярно этой поверхности
Активность нуклида (радиоактивного вещества)	$\text{Т}^{-1}$	беккерель	Бк	Bq	Беккерель есть активность нуклида в радиоактивном источнике, при которой за время $1 \text{ с}$ происходит один акт распада
Удельная активность радиоактивного вещества	$\text{М}^{-1} \text{Т}^{-1}$	беккерель на килограмм	$\text{Бк}/\text{кг}$	$\text{Bq}/\text{kg}$	Беккерель на килограмм есть удельная активность, при которой на массу $1 \text{ кг}$ радиоактивного вещества приходится активность, равная $1 \text{ Бк}$
Объемная активность радиоактивного вещества	$\text{Л}^3 \text{Т}^{-1}$	беккерель на метр кубический	$\text{Бк}/\text{м}^3$	$\text{Bq}/\text{m}^3$	Беккерель на метр кубический есть объемная активность, при которой на объем $1 \text{ м}^3$ , занимаемый радиоактивным веществом, приходится активность $1 \text{ Бк}$
Молярная активность радиоактивного вещества	$\text{Т}^{-1} \text{М}^{-1}$	беккерель на моль	$\text{Бк}/\text{моль}$	$\text{Bq}/\text{mol}$	Беккерель на моль есть молярная активность, при которой на количество вещества радионуклида $1 \text{ моль}$ приходится активность $1 \text{ Бк}$



## Приложение 2. Множители и приставки к единицам системы СИ

Множитель	Наименование (приставка)	Обозначение	
		Русское	Международное
$10^{24}$	йотта	И	Y
$10^{21}$	зетта	З	Z
$10^{18}$	экса	Э	E
$10^{15}$	пета	П	P
$10^{12}$	тера	Т	T
$10^9$	гига	Г	G
$10^6$	мега	М	M
$10^3$	кило	к	k
$10^2$	гекто	г	h
$10^1$	дека	да	da
$10^{-1}$	деци	д	d
$10^{-2}$	санتي	с	c
$10^{-3}$	милли	м	m
$10^{-6}$	микро	мк	$\mu$
$10^{-9}$	нано	н	n
$10^{-12}$	пико	п	p
$10^{-15}$	фемто	ф	f
$10^{-18}$	атто	а	a
$10^{-21}$	zepto	з	z
$10^{-24}$	йокто	и	y

### Приложение 3. Критические значения $\chi^2$ -распределения (Пирсона)

$f \backslash P$	0,01	0,05	0,10	0,50	0,90	0,95	0,99
1	0,000157	0,00393	0,0158	0,455	2,71	3,84	6,64
2	0,0201	0,103	0,211	1,39	4,61	5,99	9,21
3	0,115	0,352	0,584	2,37	6,25	7,81	11,3
4	0,297	0,711	1,06	3,36	7,78	9,49	13,3
5	0,554	1,15	1,61	4,35	9,24	11,1	15,1
6	0,872	1,64	2,20	5,35	10,6	12,6	16,8
7	1,24	2,17	2,83	6,35	12,0	14,1	18,5
8	1,65	2,73	3,49	7,34	13,4	15,5	20,1
9	2,09	3,33	4,17	8,34	14,7	16,9	21,7
10	2,56	3,94	4,87	9,34	16,0	18,3	23,2
11	3,05	4,57	5,58	10,3	17,3	19,7	24,7
12	3,57	5,23	6,30	11,3	18,5	21,0	26,2
13	4,11	5,89	7,04	12,3	19,8	22,4	27,7
14	4,66	6,57	7,79	13,3	21,1	23,7	29,1
15	5,23	7,26	8,55	14,3	22,3	25,0	30,6
16	5,81	7,96	9,31	15,3	23,5	26,3	32,0
17	6,41	8,67	10,1	16,3	24,8	27,6	33,4
18	7,01	9,39	10,9	17,3	26,0	28,9	34,8
19	7,63	10,1	11,7	18,3	27,2	30,1	36,2
20	8,26	10,9	12,4	19,3	28,4	31,4	37,6
21	8,90	11,6	13,2	20,3	29,6	32,7	38,9
22	9,54	12,3	14,0	21,3	30,8	33,9	40,3
23	10,2	13,1	14,8	22,3	32,0	35,2	41,6
24	10,9	13,8	15,7	23,3	33,2	36,4	43,0
25	11,5	14,6	16,5	24,3	34,4	37,7	44,3
30	15,0	18,5	20,6	29,3	40,3	43,8	50,9
35	18,5	22,5	24,8	34,3	47,0	49,8	57,3
40	22,2	26,5	29,0	39,3	51,8	55,8	63,7
45	25,9	30,6	33,4	44,3	57,6	61,7	69,6
50	29,7	34,8	37,7	49,3	63,2	67,6	76,2
60	37,5	43,2	46,5	59,3	74,4	79,1	88,4
70	45,4	51,7	55,3	69,3	85,5	90,5	100
80	53,5	62,4	64,3	79,3	96,6	102	112
90	61,8	69,1	73,3	89,3	108	113	124
100	70,1	77,9	82,4	99,3	118	124	136

Примечание. Более полные таблицы критических значений  $\chi^2$ -распределения см. в [17].

## Приложение 4. t-распределение Стьюдента

Критические значения для нормированного распределения Стьюдента в зависимости от числа степеней свободы и доверительной вероятности  $\bar{P}$  для одностороннего критерия.

Вероятность  $P\{|t| < |t(\bar{P}, f)|\}$

$f \backslash \bar{P}$	0,40	0,80	0,90	0,95	0,975	0,99	0,995	0,9995
1	0,325	1,376	3,078	6,314	12,71	31,82	63,65	636,6
2	0,289	1,061	1,886	2,920	4,303	6,965	9,925	31,60
3	0,277	0,978	1,638	2,353	3,182	4,541	5,841	12,94
4	0,271	0,941	1,533	2,132	2,776	3,747	4,604	8,610
5	0,267	0,920	1,476	2,016	2,571	3,365	4,032	6,859
6	0,265	0,906	1,440	1,943	2,447	3,143	3,707	5,969
7	0,263	0,896	1,415	1,895	2,365	2,998	3,499	5,405
8	0,262	0,889	1,397	1,860	2,306	2,896	3,355	5,041
9	0,261	0,883	1,383	1,833	2,262	2,821	3,250	4,781
10	0,260	0,879	1,372	1,812	2,228	2,764	3,169	4,587
11	0,260	0,876	1,363	1,796	2,201	2,718	3,106	4,437
12	0,259	0,873	1,356	1,782	2,179	2,681	3,055	4,318
13	0,259	0,870	1,350	1,771	2,160	2,650	3,012	4,221
14	0,258	0,863	1,345	1,761	2,145	2,624	2,977	4,140
15	0,258	0,866	1,341	1,753	2,131	2,602	2,947	4,073
16	0,258	0,865	1,337	1,746	2,120	2,583	2,921	4,015
17	0,257	0,863	1,333	1,740	2,110	2,567	2,898	3,965
18	0,257	0,862	1,330	1,734	2,101	2,552	2,878	3,922
19	0,257	0,861	1,328	1,729	2,093	2,539	2,861	3,883
20	0,257	0,860	1,325	1,725	2,086	2,528	2,845	3,850
21	0,257	0,859	1,323	1,721	2,080	2,518	2,831	3,819
22	0,256	0,858	1,321	1,717	2,074	2,508	2,819	3,792
23	0,256	0,858	1,319	1,714	2,069	2,500	2,807	3,767
24	0,256	0,857	1,318	1,711	2,064	2,492	2,797	3,745
25	0,256	0,856	1,316	1,708	2,060	2,485	2,787	3,725
26	0,256	0,856	1,315	1,706	2,056	2,479	2,779	3,707
27	0,256	0,855	1,314	1,703	2,052	2,473	2,771	3,690
28	0,256	0,855	1,313	1,701	2,048	2,467	2,763	3,674
29	0,256	0,854	1,311	1,699	2,045	2,462	2,756	3,659
30	0,256	0,854	1,310	1,697	2,042	2,457	2,750	3,646
40	0,255	0,851	1,303	1,684	2,021	2,423	2,704	3,551
60	0,254	0,848	1,296	1,671	2,000	2,390	2,660	3,460
120	0,254	0,845	1,289	1,658	1,980	2,358	2,617	3,373
$\infty$	0,253	0,842	1,282	1,645	1,960	2,326	2,576	3,291



# Приложение 5. F-распределение (Фишера)

В таблице приведены значения процентных точек  $F$ -распределения (пределы интегрирования этого распределения в зависимости от чисел степеней свободы  $f_1$  и  $f_2$ ) для двух уровней доверительной вероятности:  $P = 0,95$  (курсив) и  $P = 0,99$  (жирный шрифт)  $f_1$  – число степеней свободы для большей дисперсии;  $f_2$  – число степеней свободы для меньшей дисперсии.

$f_1 \backslash f_2$	1	2	3	4	5	6	7	8	9	10	12	14	16	20	24	30	40	50	100	200	$\infty$
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,41	19,42	19,43	19,44	19,45	19,46	19,47	19,47	19,49	19,49	19,50
	<b>98,49</b>	<b>99,00</b>	<b>99,17</b>	<b>99,25</b>	<b>99,30</b>	<b>99,33</b>	<b>99,34</b>	<b>99,36</b>	<b>99,38</b>	<b>99,40</b>	<b>99,42</b>	<b>99,43</b>	<b>99,44</b>	<b>99,45</b>	<b>99,46</b>	<b>99,47</b>	<b>99,48</b>	<b>99,48</b>	<b>99,49</b>	<b>99,49</b>	<b>99,50</b>
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,74	8,71	8,69	8,66	8,64	8,62	8,60	8,58	8,56	8,54	8,53
	<b>34,12</b>	<b>30,82</b>	<b>29,46</b>	<b>28,71</b>	<b>28,24</b>	<b>27,91</b>	<b>27,67</b>	<b>27,49</b>	<b>27,34</b>	<b>27,23</b>	<b>27,05</b>	<b>26,92</b>	<b>26,83</b>	<b>26,69</b>	<b>26,60</b>	<b>26,50</b>	<b>26,41</b>	<b>26,35</b>	<b>26,23</b>	<b>26,18</b>	<b>26,12</b>
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,87	5,84	5,80	5,77	5,74	5,71	5,70	5,66	5,65	5,63
	<b>21,20</b>	<b>18,00</b>	<b>16,69</b>	<b>15,93</b>	<b>15,52</b>	<b>15,21</b>	<b>14,93</b>	<b>14,80</b>	<b>14,66</b>	<b>14,54</b>	<b>14,37</b>	<b>14,24</b>	<b>14,15</b>	<b>14,02</b>	<b>13,93</b>	<b>13,83</b>	<b>13,74</b>	<b>13,69</b>	<b>13,57</b>	<b>13,52</b>	<b>13,46</b>
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,68	4,64	4,60	4,56	4,53	4,50	4,46	4,44	4,40	4,38	4,36
	<b>16,26</b>	<b>13,27</b>	<b>12,06</b>	<b>11,39</b>	<b>10,97</b>	<b>10,67</b>	<b>10,45</b>	<b>10,27</b>	<b>10,15</b>	<b>10,05</b>	<b>9,89</b>	<b>9,77</b>	<b>9,68</b>	<b>9,55</b>	<b>9,47</b>	<b>9,38</b>	<b>9,29</b>	<b>9,24</b>	<b>9,13</b>	<b>9,07</b>	<b>9,02</b>
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,71	3,69	3,67
	<b>13,74</b>	<b>10,92</b>	<b>9,78</b>	<b>9,15</b>	<b>8,75</b>	<b>8,47</b>	<b>8,26</b>	<b>8,10</b>	<b>7,98</b>	<b>7,87</b>	<b>7,72</b>	<b>7,60</b>	<b>7,52</b>	<b>7,39</b>	<b>7,31</b>	<b>7,23</b>	<b>7,14</b>	<b>7,09</b>	<b>6,99</b>	<b>6,94</b>	<b>6,88</b>
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,57	3,52	3,49	3,44	3,41	3,38	3,34	3,32	3,28	3,25	3,23
	<b>12,25</b>	<b>9,55</b>	<b>8,45</b>	<b>7,85</b>	<b>7,46</b>	<b>7,19</b>	<b>7,00</b>	<b>6,84</b>	<b>6,71</b>	<b>6,62</b>	<b>6,47</b>	<b>6,35</b>	<b>6,27</b>	<b>6,15</b>	<b>6,07</b>	<b>5,99</b>	<b>5,91</b>	<b>5,85</b>	<b>5,75</b>	<b>5,70</b>	<b>5,65</b>
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,28	3,23	3,20	3,15	3,12	3,08	3,05	3,03	2,98	2,96	2,93
	<b>11,26</b>	<b>8,65</b>	<b>7,59</b>	<b>7,01</b>	<b>6,63</b>	<b>6,37</b>	<b>6,19</b>	<b>6,03</b>	<b>5,91</b>	<b>5,82</b>	<b>5,67</b>	<b>5,56</b>	<b>5,48</b>	<b>5,36</b>	<b>5,28</b>	<b>5,20</b>	<b>5,11</b>	<b>5,06</b>	<b>4,96</b>	<b>4,91</b>	<b>4,86</b>
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,07	3,02	2,98	2,93	2,90	2,86	2,82	2,80	2,76	2,73	2,71
	<b>10,56</b>	<b>8,02</b>	<b>6,99</b>	<b>6,42</b>	<b>6,06</b>	<b>5,80</b>	<b>5,62</b>	<b>5,47</b>	<b>5,35</b>	<b>5,26</b>	<b>5,11</b>	<b>5,00</b>	<b>4,92</b>	<b>4,80</b>	<b>4,73</b>	<b>4,64</b>	<b>4,56</b>	<b>4,51</b>	<b>4,41</b>	<b>4,36</b>	<b>4,31</b>
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,91	2,86	2,82	2,77	2,74	2,70	2,67	2,64	2,59	2,56	2,54
	<b>10,04</b>	<b>7,56</b>	<b>6,55</b>	<b>5,99</b>	<b>5,64</b>	<b>5,39</b>	<b>5,21</b>	<b>5,06</b>	<b>4,95</b>	<b>4,85</b>	<b>4,71</b>	<b>4,60</b>	<b>4,52</b>	<b>4,41</b>	<b>4,33</b>	<b>4,25</b>	<b>4,17</b>	<b>4,12</b>	<b>4,01</b>	<b>3,96</b>	<b>3,91</b>
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,79	2,74	2,70	2,65	2,61	2,57	2,53	2,50	2,45	2,42	2,40
	<b>9,65</b>	<b>7,20</b>	<b>6,22</b>	<b>5,67</b>	<b>5,32</b>	<b>5,07</b>	<b>4,88</b>	<b>4,74</b>	<b>4,63</b>	<b>4,54</b>	<b>4,40</b>	<b>4,29</b>	<b>4,21</b>	<b>4,10</b>	<b>4,02</b>	<b>3,94</b>	<b>3,86</b>	<b>3,80</b>	<b>3,70</b>	<b>3,66</b>	<b>3,60</b>
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,69	2,64	2,60	2,54	2,50	2,46	2,42	2,40	2,35	2,32	2,30
	<b>9,33</b>	<b>6,93</b>	<b>5,95</b>	<b>5,41</b>	<b>5,06</b>	<b>4,82</b>	<b>4,65</b>	<b>4,50</b>	<b>4,39</b>	<b>4,30</b>	<b>4,16</b>	<b>4,05</b>	<b>3,98</b>	<b>3,86</b>	<b>3,78</b>	<b>3,70</b>	<b>3,61</b>	<b>3,56</b>	<b>3,46</b>	<b>3,41</b>	<b>3,36</b>

Приложение 5. F-распределение (Фишера) (продолжение)

$\frac{f_1}{f_2}$	1	2	3	4	5	6	7	8	9	10	12	14	16	20	24	30	40	50	100	200	$\infty$
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,60	2,55	2,51	2,46	2,42	2,38	2,34	2,32	2,26	2,24	2,21
14	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,85	3,78	3,67	3,59	3,51	3,42	3,37	3,27	3,21	3,16
15	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,53	2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,19	2,16	2,13
16	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,70	3,62	3,51	3,43	3,34	3,26	3,21	3,11	3,06	3,00
17	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,48	2,43	2,39	2,33	2,29	2,25	2,21	2,18	2,12	2,10	2,07
18	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,56	3,48	3,36	3,29	3,20	3,12	3,07	2,97	2,92	2,87
19	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,37	2,33	2,28	2,24	2,20	2,16	2,13	2,07	2,04	2,01
20	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,45	3,37	3,25	3,18	3,10	3,01	2,96	2,86	2,80	2,75
21	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,38	2,33	2,29	2,23	2,19	2,15	2,11	2,08	2,02	1,99	1,96
22	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,63	3,59	3,45	3,35	3,27	3,16	3,08	3,00	2,92	2,86	2,76	2,70	2,65
23	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,29	2,25	2,19	2,15	2,11	2,07	2,04	1,98	1,95	1,92
24	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,37	3,27	3,19	3,07	3,00	2,91	2,83	2,78	2,68	2,62	2,57
25	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,31	2,26	2,21	2,15	2,11	2,07	2,02	2,00	1,94	1,91	1,83
26	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,19	3,12	3,00	2,92	2,84	2,76	2,70	2,60	2,54	2,49
27	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,28	2,23	2,18	2,12	2,08	2,04	1,99	1,96	1,90	1,87	1,84
28	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	8,37	3,23	3,13	3,05	2,94	2,86	2,77	2,69	2,63	2,53	2,47	2,42
29	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,20	2,15	2,09	2,05	2,00	1,96	1,93	1,87	1,84	1,81
30	8,02	5,78	4,87	4,37	4,04	3,81	3,65	3,51	3,40	3,31	3,17	3,07	2,99	2,88	2,80	2,72	2,63	2,58	2,47	2,42	2,36
31	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,23	2,18	2,13	2,07	2,03	1,98	1,93	1,91	1,84	1,81	1,78
32	7,94	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	3,02	2,94	2,83	2,75	2,67	2,58	2,53	2,42	2,37	2,31
33	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,20	2,14	2,10	2,04	2,00	1,96	1,91	1,88	1,82	1,79	1,76
34	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,97	2,89	2,78	2,70	2,62	2,53	2,43	2,37	2,32	2,26
35	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,18	2,13	2,09	2,02	1,98	1,94	1,89	1,86	1,80	1,76	1,73
36	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,25	3,17	3,03	2,93	2,85	2,74	2,66	2,58	2,49	2,44	2,33	2,27	2,21
37	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,16	2,11	2,06	2,00	1,96	1,92	1,87	1,84	1,77	1,74	1,71
38	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,21	3,13	2,99	2,89	2,81	2,70	2,62	2,54	2,45	2,40	2,29	2,23	2,17



## Приложение 5. F-распределение (Фишера) (продолжение)

$\frac{f_1}{f_2}$	1	2	3	4	5	6	7	8	9	10	12	14	16	20	24	30	40	50	100	200	$\infty$
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,10	2,05	1,99	1,95	1,90	1,85	1,82	1,76	1,72	1,69
27	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,17	3,09	2,96	2,86	2,77	2,66	2,53	2,50	2,41	2,36	2,25	2,19	2,13
28	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,13	2,08	2,03	1,97	1,93	1,88	1,84	1,80	1,74	1,71	1,67
29	7,68	5,49	4,60	4,11	3,79	3,56	3,39	3,26	3,14	3,06	2,93	2,83	2,74	2,63	2,55	2,47	2,38	2,33	2,21	2,16	2,10
30	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,12	2,06	2,02	1,96	1,91	1,87	1,81	1,78	1,72	1,69	1,65
31	7,64	5,45	4,57	4,07	3,76	3,53	3,36	3,23	3,11	3,03	2,90	2,80	2,71	2,60	2,52	2,44	2,35	2,30	2,18	2,13	2,06
32	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18	2,10	2,05	2,00	1,94	1,90	1,85	1,80	1,77	1,71	1,68	1,64
33	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,08	3,00	2,87	2,77	2,68	2,57	2,49	2,41	2,32	2,27	2,15	2,10	2,03
34	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,09	2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,69	1,66	1,62
35	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,06	2,98	2,84	2,74	2,66	2,55	2,47	2,38	2,29	2,24	2,13	2,07	2,01
36	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,07	2,02	1,97	1,91	1,86	1,82	1,76	1,74	1,67	1,64	1,59
37	7,50	5,34	4,46	3,97	3,66	3,42	3,25	3,12	3,01	2,94	2,80	2,70	2,62	2,51	2,42	2,34	2,25	2,20	2,03	2,02	1,96
38	4,13	3,28	2,88	2,65	2,49	2,38	2,30	2,23	2,17	2,12	2,05	2,00	1,95	1,89	1,84	1,80	1,74	1,71	1,64	1,61	1,57
39	7,44	5,29	4,42	3,93	3,61	3,38	3,21	3,08	2,97	2,89	2,76	2,66	2,58	2,47	2,38	2,30	2,21	2,15	2,04	1,98	1,91
40	4,11	3,26	2,86	2,63	2,48	2,36	2,28	2,21	2,15	2,10	2,03	1,98	1,93	1,87	1,82	1,78	1,72	1,69	1,62	1,59	1,55
41	7,39	5,25	4,38	3,89	3,58	3,35	3,18	3,04	2,94	2,86	2,72	2,62	2,54	2,43	2,35	2,26	2,17	2,12	2,00	1,94	1,87
42	4,10	3,25	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,02	1,96	1,92	1,85	1,80	1,76	1,71	1,67	1,60	1,57	1,53
43	7,35	5,21	4,34	3,86	3,54	3,32	3,15	3,02	2,91	2,82	2,69	2,59	2,51	2,40	2,32	2,22	2,14	2,08	1,97	1,90	1,84
44	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,00	1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,59	1,55	1,51
45	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,88	2,80	2,66	2,56	2,49	2,37	2,29	2,20	2,11	2,05	1,94	1,88	1,81
46	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	1,99	1,94	1,89	1,82	1,78	1,73	1,68	1,64	1,57	1,54	1,49
47	7,27	5,15	4,29	3,80	3,49	3,26	3,10	2,96	2,86	2,77	2,64	2,54	2,46	2,35	2,26	2,17	2,08	2,02	1,91	1,85	1,78
48	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	1,98	1,92	1,88	1,81	1,76	1,72	1,66	1,63	1,56	1,52	1,48
49	7,24	5,12	4,26	3,78	3,46	3,24	3,07	2,94	2,84	2,75	2,62	2,52	2,44	2,32	2,24	2,15	2,06	2,00	1,88	1,82	1,75
50	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,14	2,09	2,04	1,97	1,91	1,87	1,80	1,75	1,71	1,65	1,62	1,54	1,51	1,46
51	7,21	5,10	4,24	3,76	3,44	3,22	3,05	2,92	2,82	2,73	2,60	2,50	2,42	2,30	2,22	2,13	2,04	1,98	1,86	1,80	1,72



Приложение 5. F-распределение (Фишера) (окончание)

$\frac{f_1}{f_2}$	1	2	3	4	5	6	7	8	9	10	12	14	16	20	24	30	40	50	100	200	$\infty$
48	4,04	3,19	2,80	2,56	2,41	2,30	2,21	2,14	2,08	2,03	1,96	1,90	1,86	1,79	1,74	1,70	1,64	1,61	1,53	1,50	1,45
50	7,19	5,08	4,22	3,74	3,42	3,20	3,04	2,90	2,80	2,71	2,58	2,48	2,40	2,28	2,20	2,11	2,02	1,96	1,84	1,78	1,70
	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,95	1,90	1,85	1,78	1,74	1,69	1,63	1,60	1,52	1,48	1,44
55	7,17	5,06	4,20	3,72	3,41	3,18	3,02	2,88	2,78	2,70	2,56	2,46	2,39	2,26	2,13	2,10	2,00	1,94	1,82	1,76	1,63
	4,02	3,17	2,78	2,54	2,38	2,27	2,18	2,11	2,05	2,00	1,93	1,88	1,83	1,76	1,72	1,67	1,61	1,58	1,50	1,46	1,41
60	7,12	5,01	4,16	3,68	3,37	3,15	2,98	2,85	2,75	2,66	2,53	2,43	2,35	2,23	2,15	2,06	1,96	1,90	1,78	1,71	1,64
	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,86	1,81	1,75	1,70	1,65	1,59	1,56	1,48	1,44	1,39
65	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,40	2,32	2,20	2,12	2,03	1,93	1,87	1,74	1,68	1,60
	3,99	3,14	2,75	2,51	2,36	2,24	2,15	2,08	2,02	1,98	1,90	1,85	1,80	1,73	1,68	1,63	1,57	1,54	1,46	1,42	1,37
70	7,04	4,95	4,10	3,62	3,31	3,09	2,93	2,79	2,70	2,61	2,47	2,37	2,30	2,18	2,09	2,00	1,90	1,84	1,71	1,64	1,56
	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,89	1,84	1,79	1,72	1,67	1,62	1,56	1,53	1,45	1,40	1,35
80	7,01	4,92	4,08	3,60	3,29	3,07	2,91	2,77	2,67	2,59	2,45	2,35	2,28	2,15	2,07	1,93	1,88	1,82	1,69	1,62	1,53
	3,96	3,11	2,72	2,48	2,33	2,21	2,12	2,05	1,99	1,95	1,88	1,82	1,77	1,70	1,65	1,60	1,54	1,51	1,42	1,38	1,32
100	6,96	4,88	4,04	3,56	3,25	3,04	2,87	2,74	2,64	2,55	2,41	2,32	2,24	2,11	2,03	1,94	1,84	1,78	1,65	1,57	1,49
	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,85	1,79	1,75	1,68	1,63	1,57	1,51	1,48	1,39	1,34	1,28
125	6,90	4,82	3,98	3,51	3,20	2,99	2,82	2,69	2,59	2,51	2,36	2,26	2,19	2,06	1,98	1,89	1,79	1,73	1,59	1,51	1,43
	3,92	3,07	2,68	2,44	2,29	2,17	2,08	2,01	1,95	1,90	1,83	1,77	1,72	1,65	1,60	1,55	1,49	1,45	1,36	1,31	1,25
150	6,84	4,78	3,94	3,47	3,17	2,95	2,79	2,65	2,56	2,47	2,33	2,23	2,15	2,03	1,94	1,85	1,75	1,68	1,54	1,46	1,37
	3,91	3,06	2,67	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,82	1,76	1,71	1,64	1,59	1,54	1,47	1,44	1,34	1,29	1,22
200	6,81	4,75	3,91	3,44	3,14	2,92	2,76	2,62	2,53	2,44	2,30	2,20	2,12	2,00	1,91	1,83	1,72	1,66	1,51	1,43	1,33
	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,80	1,74	1,69	1,62	1,57	1,52	1,45	1,42	1,32	1,26	1,19
400	6,76	4,71	3,88	3,41	3,11	2,90	2,73	2,60	2,50	2,41	2,23	2,17	2,09	1,97	1,88	1,79	1,69	1,62	1,48	1,39	1,28
	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,78	1,72	1,67	1,60	1,54	1,49	1,42	1,38	1,28	1,22	1,13
∞	6,70	4,66	3,83	3,36	3,06	2,85	2,69	2,55	2,46	2,37	2,23	2,12	2,04	1,92	1,84	1,74	1,64	1,57	1,42	1,32	1,19
	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,75	1,69	1,64	1,57	1,52	1,46	1,40	1,35	1,24	1,17	1,00
	6,63	4,60	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,07	1,99	1,87	1,79	1,69	1,59	1,52	1,36	1,25	1,00

# Приложение 6. Критические значения теста Диксона $Q_{i, k \text{ крит}}$ (для $N = 3-20$ )

Статистическая функция $Q_{i,k}$	Число измерений $N$	Степень риска (вероятность ошибки 2-го рода)			
		0,005	0,01	0,05	0,10
$Q_{1,0}$	3	0,994	0,988	0,941	0,886
	4	0,926	0,889	0,765	0,679
	5	0,821	0,780	0,642	0,557
	6	0,740	0,698	0,560	0,482
	7	0,680	0,637	0,507	0,434
$Q_{1,1}$	8	0,725	0,683	0,554	0,479
	9	0,677	0,635	0,512	0,441
	10	0,639	0,597	0,477	0,409
$Q_{2,1}$	11	0,713	0,679	0,576	0,517
	12	0,675	0,642	0,546	0,490
	12	0,649	0,615	0,521	0,467
$Q_{2,2}$	14	0,674	0,641	0,546	0,492
	15	0,647	0,616	0,525	0,472
	16	0,624	0,595	0,507	0,454
	17	0,605	0,577	0,490	0,438
	18	0,589	0,561	0,475	0,424
	19	0,575	0,547	0,462	0,412
	20	0,562	0,535	0,450	0,401

## Приложение 7. Непараметрический критерий $U$ (Вилкоксона – Манна – Уитни)

Максимальное число инверсий  $U(\bar{P}, N_1, N_2)$  при котором различия между группами наблюдений можно считать значимыми с  $P=0,95$ . В таблице даны значения для  $N_1, N_2 < 26$ .\*

$N_1 \backslash N_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
2	—																							
3	—	0																						
4	—	0	1																					
5	0	1	2	4																				
6	0	2	3	5	7																			
7	0	2	4	6	8	11																		
8	1	3	5	8	10	13	15																	
9	1	4	6	9	12	15	18	21																
10	1	4	7	11	14	17	20	24	27															
11	1	5	8	12	16	19	23	27	31	34														
12	2	5	9	13	17	21	26	30	34	38	42													
13	2	6	10	15	19	24	28	33	37	42	47	51												
14	3	7	11	16	21	26	31	36	41	46	51	56	61											
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72										
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83									
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96								
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109							
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123						
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138					
21			19	26	34	41	49	57	65	73	81	89	97	105	113	121	130	138	146	154				
22			20	28	36	44	52	60	69	77	85	94	102	111	119	128	136	145	154	162	171			
23			21	29	37	46	55	63	72	81	90	99	107	116	125	134	143	152	161	170	180	189		
24			22	31	39	48	57	66	75	85	94	103	113	122	131	141	150	160	169	179	188	198	207	
25			23	33	41	50	60	69	79	89	98	108	118	128	137	147	157	167	177	187	197	207	217	227

\* Значения  $U(\bar{P}, N_1, N_2)$  для больших  $N$  приведены в [17].



## Приложение 8. Критерий знаков

Максимальное число наименее часто встречающихся знаков  $N_{кр}$ , при котором гипотеза, что две связанные выборки получены из одной совокупности, отвергается (двухсторонний критерий) с доверительными вероятностями  $\bar{P} = 0,95$  и  $\bar{P} = 0,99$ .

$m$	$\bar{P}$		$m$	$\bar{P}$		$m$	$\bar{P}$		$m$	$\bar{P}$		$m$	$\bar{P}$	
	0,95	0,99		0,95	0,99		0,95	0,99		0,95	0,99		0,95	0,99
5	0	—	22	6	5	39	13	11	62	24	21	96	39	36
6	0	—	23	7	5	40	14	12	64	24	22	98	40	37
7	0	0	24	7	5	41	14	12	66	25	23	100	41	37
8	1	0	25	7	6	42	15	13	68	26	23	110	45	42
9	1	0	26	8	6	43	15	13	70	27	24	120	50	46
10	1	0	27	8	7	44	16	13	72	28	25	130	55	51
11	2	1	28	8	7	45	16	14	74	29	26	140	59	55
12	2	1	29	9	7	46	16	14	76	30	27	150	64	60
13	3	1	30	10	8	47	17	15	78	31	28	160	69	64
14	3	2	31	10	8	48	17	15	80	32	29	170	73	69
15	3	2	32	10	8	49	18	15	82	33	30	180	78	73
16	4	2	33	11	9	50	18	16	84	33	30	190	83	78
17	4	3	34	11	9	52	19	17	86	34	31	200	87	83
18	5	3	35	12	10	54	20	18	88	35	32	220	97	92
19	5	4	36	12	10	56	21	18	90	36	33	240	106	101
20	5	4	37	13	10	58	22	19	92	37	34	260	116	110
21	6	4	38	13	11	60	23	20	94	38	35	280	125	120
												300	135	129

## Приложение 9. Парный критерий Вилкоксона

Величины  $N_{\text{крит}}$  (максимальные значения  $T$ , при которых гипотеза, что две связанные выборки получены из одной совокупности, отвергается (двухсторонний критерий) с доверительными вероятностями  $\bar{P} = 0,95$  и  $\bar{P} = 0,99$ .

$N$	$\bar{P}$		$N$	$\bar{P}$		$N$	$\bar{P}$	
	0,95	0,99		0,95	0,99		0,95	0,99
5	0	—	11	13	7	17	41	28
6	2	0	12	17	10	18	47	33
7	3	0	13	21	12	19	53	38
8	5	1	14	25	16	20	60	42
9	8	3	15	30	19			
10	10	5	16	35	23			

## Приложение 10

Минимальные значения абсолютных значений коэффициентов ранговой корреляции (Спирмена)  $|R|$ , при которых гипотезу об отсутствии связи между двумя рядами результатов измерений можно отвергнуть с доверительной вероятностью  $\bar{P}$  ( $4 \leq m \leq 10$ ) [17].

$m$	Доверительная вероятность $\bar{P}$				
	0,05	0,025	0,01	0,005	0,001
4	1,000				
5	0,900	1,000	1,000	1,000	
6	0,771	0,828	0,886	0,942	1,000
7	0,678	0,768	0,836	0,893	0,964
8	0,643	0,714	0,786	0,857	0,928
9	0,633	0,700	0,767	0,833	0,900
10	0,564	0,685	0,746	0,806	0,867

## Приложение 11. Краткое описание Федеральной системы внешней оценки качества клинического анализа

Необходимость внешней оценки качества анализа для клинических лабораторий общепризнанна. В ряду других мероприятий по обеспечению качества анализа внешняя оценка качества занимает особое место, поскольку обеспечивает объективную оценку качества исследований по конечному результату. Надо отметить, что требования к качеству анализа в медицине особенно жесткие из-за непосредственного влияния результатов анализа на здоровье людей. Кроме того, экономические ограничения заставляют проводить только одно измерение для каждой пробы и системы обеспечения и контроля качества анализа являются единственной гарантией надежности получаемых результатов.

К 1990–91 гг. была осознана необходимость создания национальной системы внешней оценки качества. При этом учитывались известные преимущества крупных систем:

- возможность привлечения квалифицированных специалистов;
- возможность надежной аттестации материалов сравнения внутри системы из-за больших ее размеров;
- низкая стоимость участия в системе в расчете на одну лабораторию;
- возможность изучения состояния и общих тенденций клинического анализа в стране, и так далее.

В 1994 году Министерство здравоохранения России определило конкретные меры по построению полномасштабной системы внешней оценки качества, охватывающей все основные разделы лабораторной медицины, в том числе методы химического анализа. Были созданы Центр внешнего контроля качества клинических лабораторных исследований (на базе Отдела стандартизации и контроля качества лабораторных исследований Национального центра для профилактической медицины) и экспертные группы по разработке методов внешней оценки качества для конкретных дисциплин лабораторной медицины. В 1995 г. начала работать федеральная система внешней оценки качества России (ФСВОК). Она функционирует без финансовой поддержки Министерства здравоохранения России, но под его контролем: расходы покрываются лабораториями-участниками системы.

**Основной декларируемой целью ФСВОК** является помощь клиническим лабораториям в объективной оценке качества выполняемых исследований и выработке рекомендаций по его повышению. Наряду с оценкой качества анализа и выдачей конкретных рекомендаций для лабораторий, ФСВОК информирует лаборатории и органы управления здравоохранением о сравнительном качестве используемых в стране наборов реактивов, калибровочных материалов и оборудования, а также о новых средствах и методах анализа.

Для достижения этих целей в ФСВОК много усилий направляется на обеспечение условий, в которых лаборатории не видели бы необходимости в создании особых условий при исследовании получаемых материалов сравнения. Принимаются меры по недопущению каких-либо административных санкций по результатам оценки качества анализов, что обеспечивается:



- анонимностью лабораторий: все лаборатории кодируются, информация о качестве исследований в каждой лаборатории рассматривается как конфиденциальная и сообщается только руководителю лаборатории;
- тем, что для аккредитации лаборатории необходимо лишь участие в ФСВОК, а не выполнение ее критериев. Такое положение будет сохранено еще по крайней мере два года.

Кроме того, объективность оценок обеспечивается:

- кодированием каждого флакона с образцом сравнения;
- проведением параллельных измерений из разных флаконов;
- требованием выполнения анализов в рутинной серии исследований обычных проб, поступающих в лабораторию на анализ, в тех же условиях, с теми же реагентами, на том же оборудовании;
- просмотром полученных из лабораторий данных экспертами ФСВОК для выявления грубых ошибок (нарушения дизайна, неправильные единицы измерения и т.д.);
- дублированным вводом результатов ВОК в компьютер с последующей сверкой данных, что практически исключает возможность ошибки.

В рамках ФСВОК проводится несколько отдельных программ оценки качества, выделенных в соответствии с традициями клинической лабораторной медицины. Количественный химический анализ проводится в следующих из них: Биохимия, Анализ мочи, Определение гормонов, Гематология. Анализ результатов, полученных из лабораторий в течение 1995–97 гг., показывает систематическое повышение качества анализа в лабораториях, постоянно участвующих в ФСВОК.

**Экспериментальный дизайн, контрольные материалы и определяемые вещества.**

Исходя из целей программы, т.е. с учетом ее обучающей роли, дизайн предусматривает возможность оценки как внутри-, так и межлабораторных эффектов.

Ежегодно проводится три цикла внешней оценки качества по каждой программе. В каждом цикле в лаборатории высылаются два контрольных материала с разной концентрацией определяемых веществ. Концентрации определяемых веществ в одном из образцов лежат в диапазоне нормальных, а в другом — патологических значений. Каждая лаборатория должна выполнить по два определения каждого вещества в каждом материале. Согласно протоколу анализ выполняется одновременно во всех лабораториях.

Применяются контрольные материалы промышленного изготовления. В программах «Биохимия» и «Определение гормонов» используются лиофилизированная сыворотка крови человека, в программе «Анализ мочи» — лиофилизированная человеческая моча, в программе «Гематология» — стабилизированная цельная человеческая кровь.

Вещества, качество определения которых оценивается в программе «Биохимия», перечислены в таблице. В программе «Анализ мочи» — определение амилазы, белка, билирубина, гемоглобина, глюкозы, калия, кальция, кетоновых тел, 17-кетостероидов, креатинина, мочевины, мочевой кислоты, натрия, уробилиногена, фосфора, хлоридов. В программе «Определение гормонов» — определение концентрации альдостерона, альфа-фетопroteина, бета-ХГЧ, кортизола, лютеинизирующего гормона, РЭА, С-пептида, прогестерона, пролактина, свободных Т3 и Т4, теофиллина,

тестостерона, тиреотропного гормона, тироксина, трийодтиронина, ферритина, фоллитропина и эстрадиола. В программе «Гематология» наряду с концентрацией эритроцитов оценивается качество определения гемоглобина. Этот список ежегодно корректируется.

Таблица. Критерии внешней оценки качества. Биохимия

Определяемое вещество	Смещение единичного измерения	Смещение среднего из двух измерений	Относительный размах двух измерений
Аланинаминотрансфераза	±30%	±25%	<15%
Альбумин	±18%	±14%	<10%
Амилаза	±30%	±25%	<15%
Аспаратаминотрансфераза	±30%	±25%	<15%
Белок общий	±7%	±6%	<3%
Билирубин общий	±30%	±24%	<12%
Глюкоза	±16%	±12%	<7%
γ-глутамилтрансфераза	±30%	±25%	<15%
Железо	±30%	±23%	<13%
Калий	±10%	±8%	<5%
Кальций	±7%	±5%	<3%
Креатинин	±18%	±15%	<8%
Креатинкиназа	±30%	±25%	<15%
Лактатдегидрогеназа	±30%	±25%	<15%
Магний	±25%	±20%	<15%
Мочевая кислота	±15%	±12%	<6%
Мочевина	±20%	±16%	<7%
Натрий	±4%	±3%	<2%
Триацилглицериды	±30%	±24%	<12%
Фосфор неорганический	±13%	±10%	<6%
Хлориды	±6%	±4%	<3%
Холестерин общий	±15%	±12%	<5%
Щелочная фосфатаза	±30%	±25%	<15%

Статистическая обработка результатов КХА проводится отдельно для каждого контрольного материала.

- По результатам двух измерений  $X_1$  и  $X_2$ , полученных в каждой лаборатории, рассчитывается среднее значение  $\bar{X} = (X_1 + X_2) / 2$  и относительный размах  $R = (X_1 - X_2) / \bar{X}$ .
- Из совокупности величин  $\bar{X}$  отбрасываются по 2,5% наименьших и наибольших значений.
- Результаты группируются по используемым методам или группам методов анализа («группы»).
- Проверяется нормальность распределения величин  $\bar{X}$  в группах; при необходимости проводится нормализация.
- Рассчитываются общие средние значения  $\bar{\bar{X}} = \frac{1}{I} \sum_i \bar{X}_i$  и стандартное отклонение  $S$  в каждой группе. Величины  $\bar{X}$ , лежащие вне интервала (общее среднее  $\pm 2$  стандартных отклонения), отбрасываются.

- е. Общее среднее  $\bar{\bar{X}} = \frac{1}{J} \sum_i \bar{X}_i$  рассчитывается вновь. Полученные величины рассматриваются как целевые (в ФСВОК их называют «референтные значения», РЗ).
- ж. Рассчитываются значения относительного смещения  $B = (\bar{X} - PЗ) / PЗ$  и средние значения величин  $R(\bar{R})$ .
- з. Проверяется соответствие оцениваемых результатов (величин  $\bar{X}$  и  $R$ ) критериям программы. Это делается по разному в разделе «Биохимия» и в других программах.

В разделе «Биохимия» в качестве критериев используются фиксированные значения относительного смещения  $B$  и  $R$  (см. таблицу). Допустимыми являются величины, меньшие этих критериев. Заметим, что эти критерии утверждаются нормативным документом Минздрава России (Приказ № 45 от 07.02.2000).

В случае количественных исследований в других программах границы диапазона допустимых значений среднего и относительного размаха определяются заново в каждом из циклов ФСВОК после отброса грубых выбросов. В качестве такого диапазона для среднего используется интервал  $(PЗ \pm 1,64 \text{ стандартного отклонения})$  для распределения величин  $\bar{X}$ . С учетом 5% лабораторий, чьи результаты рассматриваются как грубые ошибки, доля лабораторий, попадающих в диапазон допустимых значений для среднего, составляет 85%. В качестве максимально допустимого относительного размаха  $R$  принимается величина, отделяющая 5% лабораторий данной группы с наибольшими значениями размаха.

Результаты оценки правильности и воспроизводимости выполненных исследований представляются каждой лаборатории отдельно для каждого определяемого вещества в графической (в виде гистограмм) и численной (таблицы) форме. Пример такого представления приведен на рис. 3.11.

На гистограммах представлено, как величины среднего и относительного размаха для каждого контрольного материала (для краткости называемого «пулом») распределились по лабораториям, участвовавшим в данном цикле ФСВОК. Часть столбцов, выделенная темно-серым цветом, соответствует лабораториям данной группы, незакрашенная часть — остальным лабораториям. Положение среднего или размаха для данной лаборатории показано стрелкой, под которой приведено его численное значение («Ваше значение»). Вертикальной линией, оканчивающейся сверху прямоугольником с меткой «РЗ», на гистограммах средних значений показаны референтные значения. Серым фоном на шкале по оси абсцисс выделены области неудовлетворительных результатов. Соответственно, интервал шкалы без фона показывает область значений величин  $\bar{X}$  и  $R$ , рассматриваемых в ФСВОК как допустимые.

При представлении результаты оценки в численном виде (в таблице) для каждого пула приводятся:

- среднее значение  $\bar{X}$ , рассчитанное для данной лаборатории,
- референтное значение РЗ (в скобках указывается, для какой совокупности лабораторий оно рассчитано),
- относительное смещение  $B$  — отклонение среднего лаборатории от референтного значения в процентах к последнему,



- диапазон допустимых значений  $\bar{X}$  для данной группы,
- число лабораторий с вашим методом (число лабораторий в группе),
- коэффициент межлабораторной вариации, % (коэффициент вариации величин  $\bar{X}$  в группе).
- среднее всех лабораторий — среднее значение для всех лабораторий, участвовавших в данном цикле, независимо от используемого метода анализа,
- общее число лабораторий, участвовавших в данном цикле, оставшееся после отброса выпадающих значений,
- коэффициент межлабораторной вариации в %, рассчитанный для всех лабораторий.

Данные для размахов приводятся отдельно:

- относительный размах измерений в данной лаборатории,
- допустимый относительный размах,
- средний относительный размах в данной группе лабораторий,
- средний относительный размах всех лабораторий.

Значения смещения и относительного размаха, выходящие за пределы диапазона допустимых значений, помечаются звездочкой (\*).

Раздельная оценка правильности и воспроизводимости позволяет сформулировать общие рекомендации, направленные на снижение конкретно систематической или случайной составляющей погрешностей. Такие рекомендации приводятся в письме, направляемом вместе с результатами оценки качества выполненных в них исследований.

По-другому оцениваются **результаты полуколичественных исследований**, широко распространенных при анализе мочи.

К полуколичественным в ФСВОК относятся методы, результат определения которыми может быть выражен одним из трех и более фиксированных заключений. При этом выделяют две группы методов: а) методы, для которых возможный результат анализа — одна из нескольких нечисловых категорий (например, «–, ±, +, ++, +++»); б) методы, для которых возможный результат анализа — один из числовых интервалов (например, для белка: «0–0,1; 0,1–0,3; 0,3–1,0; 1,0–3,0; >3,0 г/л»). Сюда же относятся методы с возможным результатом в виде числа в тех случаях, когда таких чисел немного (например, возможны лишь результаты определения глюкозы 0; 5; 17; 51; 150; более 150 ммоль/л); такие результаты для оценки преобразуются в интервалы.

#### *а) Методы с нечисловым результатом измерения*

Для таких методов выделяются группы лабораторий с одинаковым числом возможных результатов независимо от аналитического принципа используемого метода. Например, лаборатории, в которых возможен один из четырех результатов («–, +, ++, +++»), составляют одну группу. В рамках каждой группы результаты располагают в порядке уменьшения их частоты: сначала ищется результат, полученный в большинстве лабораторий. Если доля лабораторий с этим результатом меньше 85%, берется следующий по числу лабораторий результат и рассчитывается доля лабораторий с этим и предыдущим результатом и т.д. Как только доля лабораторий достигает 85%, процесс прекращается и оставшиеся результаты считаются

неудовлетворительными. Результаты оценки представляются в лаборатории в виде двух гистограмм. На гистограммах штриховкой выделяются результаты, признанные неудовлетворительными, а также показывается результат, полученный данной лабораторией.

***б) Полуколичественные методы с числовым представлением результата измерения***

В данной категории методов оценку правильности исследования производят с использованием в качестве целевого значения медианы средин всех интервалов, полученных в лабораториях. Если границы возможных интервалов расположены на числовой шкале неравномерно, проводится линеаризация. Если целевое значение попадает в числовой интервал, представленный в качестве результата измерения данной лабораторией, или в верхнюю треть предыдущего (для всех, кроме первого) интервала, или в нижние 2/3 следующего, то такой результат считается удовлетворительным, в противном случае — неудовлетворительным. Например, если концентрация белка в пуле определена лабораторией равной 1–3 г/л, предыдущий возможный интервал 0,3–1 г/л, а следующий возможный интервал не имеет верхней границы — >3 г/л, то такой результат признается удовлетворительным, если целевое значение находится в интервале от 0,77 до 7 г/л.

Оценки результатов полуколичественных измерений представляются в лаборатории в табличном виде.

**Участие в ФСВОК или других системах внешней оценки качества и полученные оценки учитываются при аккредитации клинических лабораторий.**

## Литература

1. International vocabulary of metrology – Basic and general concepts and associated terms (VIM). 3<sup>rd</sup> Ed. JCGM 200:2012. Имеется русский перевод: Международный словарь по метрологии: основные и общие понятия и соответствующие термины: пер. с англ. и фр. / Всерос. науч.-исслед. ин-т метрологии им. Д. И. Менделеева, Белорус. гос. ин-т метрологии. Изд. 2-е, испр. – СПб.: НПО «Профессионал», 2010. – 82 с.
2. Федеральный закон Российской Федерации от 26 июня 2008 г. № 102-ФЗ «Об обеспечении единства измерений».
3. Бурдун Г.Д., Марков Б.Н. Основы метрологии. – М.: Из-во стандартов, 1985. – 256 с.
4. <https://www.bipm.org/>
5. ГОСТ 8.417-2002 ГСОЕИ. Единицы величин. 2003 г. 24 с.
6. ГОСТ Р 8.563-2009. «ГСИ. Методики (методы) измерений». – М.: Стандарт-информ. 2010 г. – 26 с.
7. ГОСТ 8.610-2004. Плотность нефти (Таблицы перерасчета). – М.: ИПК Из-во стандартов, 2005.
8. Каршенбойм С.Г. // Успехи физических наук. 2006. Т. 176. № 9. С. 975–982.
9. Gläser M., Borys M. // Rep. Prog. Phys. 2009. V. 72. N. 12.
10. Robinson I.A., Schlamminger S. // Metrologia 53 (2016) A46–A74.
11. ГОСТ 8.061-80 «ГСИ. Поверочные схемы. Содержание и построение».
12. ГОСТ 8.558-2009. ГСОЕИ. Государственная поверочная схема для средств измерений температуры. – М.: Из-во стандартов. 1994. – 16 с.
13. ГОСТ 8.009-84 «ГСИ. Нормируемые метрологические характеристики средств измерений».
14. ГОСТ 8.401-80. ГСОЕИ. Классы точности средств измерений. Общие требования. М.: Из-во стандартов. 1980. 8 с.
15. ГОСТ Р 8.734-2011. ГСОЕИ. Датчики интеллектуальные и системы измерительные интеллектуальные. Методы метрологического самоконтроля.
16. Федеральный закон Российской Федерации от 27.12.2002 № 184-ФЗ «О техническом регулировании».
17. Элькин Г.И., Г.И.Саламатов Г.И., Крутиков В.Н., Новиков Н.Ю. Состояние и основные направления работ в области обеспечения единства измерений в Российской Федерации. <http://metrobr.ru>
18. Приказ Министерства промышленности и торговли РФ от 30 ноября 2009 г. № 1081.
19. Александров Ю.И. Спорные вопросы современной метрологии в химическом анализе. – С.-Петербург, 2003. – 304 с.
20. Дерффель К. Статистика в аналитической химии. – М., Мир, 1994. – 272 с.
21. Vincze I. Matematikai statisztika Ipari Alkalmazásokkal. Budapest: Muszaki Könyvkiado. 1975. 351 p.
22. Буйташ П., Кузьмин Н.М., Лейстнер Л. Обеспечение качества химического анализа. – М., Наука. 1993. – 188 с.
23. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М., Наука. 1965. – 464 с.
24. Гублер Е.В., Генкин А.А. Применение непараметрических критериев статистики в медико-биологических исследованиях. – Ленинград: Медицина, 1973. – 137 с.



25. Орлов А.И. // Заводская лаборатория. Диагностика материалов. 2012. Т. 78. № 11. С. 66–70.
26. Taylor J.K. Quality assurance of chemical measurements. Chelsea: Lewis Publ. Inc. 1988. 328 p.
27. ГОСТ Р ИСО 5725-2002 Точность (правильность и прецизионность) методов и результатов измерений. Ч. 1–6.
28. Barnett V., Lewis T. Outliers in Statistical Data. N.-Y.: Wiley. 1984. 365 p.
29. Miller J.N. // Analyst. 1993. V. 118. P. 455–461.
30. Rousseeuw P.J., Leroy A.M. Robust regression and outlier detection. N.-Y.: Wiley. 1987. 329 p.
31. Kateman G., Pijpers F.W. Quality control in analytical chemistry. N.-Y.: Wiley. 1981. Имеется русский перевод: Катеман Г., Пийперс Ф.В. Контроль качества химического анализа. Челябинск: Металлургия. 1989. 447 с.
32. Кокрен У. Методы выборочного исследования. — М.: Статистика, 1976. — 440 с.
33. Шеффе Г. Дисперсионный анализ. — М., Наука, 1980. — 512 с.
34. Deming S.N., Morgan S.L. Experimental Design: A Chemometric Approach, 2nd Rev. Amsterdam: Elsevier, in: «Data Handling in Science and Technology». V. 11, 1994. 454 p.
35. SAS/STAT User's Guide, Release 6.03 Edition. SAS Institute Inc., NC 27512-8000.
36. Theil A. // Prock. Ned. Wet. Ser. A. 1950. V. 53. P. 386.
37. Холлендер М., Вулф Д. Непараметрические методы статистики. — М.: Финансы и статистика, 1983. — 518 с.
38. Draper N.R., Smith H. Applied Regression Analysis. 3rd Ed. / NY:Wiley & Sons. 1998. 706 p.
39. Rousseeuw P.J., van Aelst S., Rambali P.J., Smeyers-Verbeke P.J. // Analytica Chimica Acta, 446 (2001), 1–2 (ноябрь 19), 243–254.
40. Parvin C.A. // Clin. Chem. 1984, V. 30. No. 5. P. 751–754.
41. Jolliffe I.T. Principal Component Analysis. Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002, XXIX, 487 p.
42. Себер Дж. Линейный регрессионный анализ. — М.: Мир, 1980. — 456 с.
43. Caulett R. Statistics in Research and Development, 2nd Ed. London etc: Chapman and Hall. 1991. 352 P.
44. Draper N.R., Smith H. Applied Regression Analysis. 3rd Ed. NY:Wiley & Sons. 1998. 706 p.
45. Seber G.A.F., Wild C.J. Nonlinear Regression. New York: John Wiley and Sons, 1989. 768 p.
46. Демиденко Е.З. Линейная и нелинейная регрессия. — М.: ФиС, 1981. — 303 с.
47. Kleinbaum D.G., Kuppel L.L. and Muller K.E. Applied Regression Analysis and other multivariable Methods, PWS-Kent, Boston, MA, 1998. XVIII, 718 p.
48. Morrison D.F. Multivariate Statistical Methods. NY: McGraw-Hill. 1990. 415 p.
49. Дворкин В.И. Метрология химического анализа. Большая российская энциклопедия. 2012. Т. 20. С. 147–148.
50. Комарь Н.П. // Ж. аналит. химии. 1978. Т. 33. Вып. 11. С. 2277.
51. Vitzthum F., Siest G., Bunk D.M. et al. // Proteomics Clin. Appl. 2007. V. 1. P. 1016–1035.
52. ГОСТ Р 8.753-2011. ГСОЕИ. Стандартные образцы материалов (веществ). Основные положения. М. Стандартинформ. 2013. 20 с.
53. Дворкин В.И., Каринова И.Н. // Ж. аналит. химии. 1990. Т.45. С. 837–857.
54. Дворкин В.И., Каринова И.Н. // Ж.аналит. химии. 1992. Т. 47. Вып. 7. С. 1200–1210.
55. Шаяевич А.Б. Стандартные образцы для аналитических целей. — М.: Химия, 1987. — 184 с.

56. ГОСТ 8.315-97 ГСИ. Стандартные образцы состава и свойств веществ и материалов (с поправкой опубл. 28.05.2002 в ИУС 11-2002).
57. *Anderson G., Scott M.* // Clin. Chem., 1991. V. 37. N. 3. P. 398–402.
58. ГСОЕИ. Стандартные образцы состава и свойств веществ и материалов. Методика оценивания характеристики стабильности. P 50.2.031-2003.
59. *Дворкин В.И., Малахов В.Н., Заикин Е.В.* // Ж. аналит. химии. 1985. Т. 40. С. 2117–2124.
60. ГОСТ Р 8.694-2010 (Руководство ИСО 35:2006) ГСИ. Стандартные образцы материалов (веществ). Общие статистические принципы определения метрологических характеристик.
61. *Дворкин В.И.* // Адекватность и неадекватность в метрологии химического анализа. Журнал аналитической химии, 2003. Т. 58. № 6. С. 566–570.
62. *Dvorkin V.* // Adequacy in the metrology of chemical analysis and other complicated measurements. Accreditation and Quality Assurance. 2004., V. 9. No. 7. P. 421–424.
63. *Дворкин В.И., Каринова И.Н.* // Ж. аналит. химии. 1991. Т. 46 № 6. С. 1045–1056.
64. *Rothman K. and Greenland S.* Types of Epidemiologic Studies. In: Rothman K., Greenland S. Modern Epidemiology. 2<sup>nd</sup> ed., Lippincott Williams & Wilkins, Philadelphia. 1998. P. 67–78.
65. *Barwick V.J., Prichard E.* (Eds), Eurachem Guide: Terminology in Analytical Measurement – Introduction to VIM 3 (2011). ISBN 978-0-948926-29-7. Имеется русский перевод: Терминология аналитических измерений. Введение в ВИМ 3 // ред. В.В. Меньшиков. – М.: Лабора, 2012. – 64 с.
66. *Kaizer H.* // Z. analyt. Chem. 1972. Bd. 260. S. 252–256.
67. РМГ 61-2010 Государственная система обеспечения единства измерений. Показатели точности, правильности, прецизионности методик количественного химического анализа. Методы оценки.
68. *Дворкин В.И.* // Журнал аналитической химии. 1999. Т. 54. № 8. С. 790–796
69. *Джонсон Н., Лион Ф.* Статистика и планирование эксперимента в науке и технике. Ч. 2. Методы планирования эксперимента. – М.: Мир, 1981. – 516 с.
70. *Wernimont G.T.* Use of statistics to develop and evaluate analytical methods. / Ed. G.T.Sprendley. Arlington (Virginia): AOAC. 1987. 183 p.
71. *Wernimont G.T.* Validation of measurement process / Ed. J.R.De Voe. Wash.: Amer. Chem. Soc. 1977. P. 1–29 (ACS Symposium Ser.: N 63).
72. *Ellison S.L.R., King B., Rösslein M., Salit M., Williams A.* Eurachem/CITAC Guide: Traceability in chemical measurement – A guide to achieving comparable results in chemical measurement, 2003.
73. ГОСТ ИСО/МЭК 17025-2009 «Общие требования к компетентности испытательных и калибровочных лабораторий». М.: Стандартиформ. См. также более позднюю версию ISO/IEC 17025:2017 «General requirements for the competence of testing and calibration laboratories».
74. Рекомендации по метрологии Р 50.2.028-2003. Алгоритмы построения градуировочных характеристик средств измерений состава веществ и материалов и оценивание их погрешностей (неопределенностей). ИПК Издательство стандартов, 2003.
75. Рекомендации по метрологии Р 50.2.038-2004. Измерения однократные прямые. Оценивание погрешностей и неопределенности результата измерений. ИПК Издательство стандартов, 2004.
76. International Vocabulary of Basic and General Terms in Metrology. ISO, Geneva, 1993. ISBN 0-948926-08-2.

77. EUROCHEM/CITAC Guide «Quantifying Uncertainty in Analytical Measurements», Second Ed., 2000. Имеется русский перевод: Руководство ЕВРАХИМ/СИТАК «Количественное описание неопределенности в аналитических измерениях», второе издание / Пер. с англ. Р.Л. Кадиса, Г.Р. Нежиховского, В.Б. Симины // ред. Л.А. Конопелько. Санкт-Петербург, 2002 г. 141 с.
78. Кузнецов В.П. // Измерительная техника. 2003. № 8. С. 21–27.
79. Cochren W.G. Sampling Techniques (2nd Ed.) J. Wiley & Sons, New York, 1963. 413 p.
80. ILAC G17: 2002 «Introducing the Concept of Uncertainty of Measurement in Testing in Association with the Application of the Standard ISO/IFS 17025». [www.ilac.org](http://www.ilac.org)
81. ISO/TS 21748:2004 «Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation». Имеется русский перевод: ГОСТ Р ИСО 21748-2012 Статистические методы. Руководство по использованию оценок повторяемости, воспроизводимости и правильности при оценке неопределенности измерений.
82. ГОСТ 26781-85 «Молоко. Метод измерения pH». — М.: Издательство стандартов, 1994.
83. ГОСТ 26931-86 «Сырье и продукты пищевые. Методы определения меди» с изм. № 1 (ИУС8-90).
84. ГОСТ Р 51930-2002 «Бензины автомобильные и авиационные. Определение бензола методом инфракрасной спектроскопии». — М., ИПК Издательство стандартов, 2002.
85. Кобцов С.Н. Разработка и метрологическое обоснование потенциометрического способа определения основного вещества в алкилметилфосфонатах. Дисс. на соискание ученой степени кандидата химических наук: Саратов, 2012. — С. 176–185.
86. Principles of Data Quality Control in Chemical Analysis / Analyst. 1989. V. 114. P. 1497–1503.
87. ASTM C802 – 09a Standard Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods for Construction Materials
88. Дворкин В.И. Заводская лаборатория, диагностика материалов. 1999. Т. 65. № 2. С. 55–63.
89. Youden W.J. Statistical manual of the Association of official analytical chemists (AOAC). Wash.: AOAC, 1975. P. 1–63.
90. Шаевич А.Б. Аналитическая служба как система. — М.: Химия, 1981. — 264 с.
91. Налимов В.В. Применение математической статистики при анализе вещества. — М.: Физматгиз, 1960. — 430 с.
92. Дворкин В.И. // Журнал аналитической химии. 1993. Т. 48. С. 1127–1142.
93. Дворкин В.И. // Chemometrics and Intelligent Laboratory Systems, 1994. V. 22. P. 127–146.
94. Чарыков А.К. Математическая обработка результатов химического анализа. — Л.: Химия, 1984. — С. 33.
95. Афифи А., Эйзен С. Статистический анализ. — Подход с использованием ЭВМ. — М.: Мир, 1982. — 488 с.
96. ГОСТ ИСО/МЭК 17011-2009. Оценка соответствия. Общие требования к органам по аккредитации, аккредитующим органы по оценке соответствия. — М.: Стандартинформ, 2010. — 24 с.
97. Федеральный закон Российской Федерации от 28 декабря 2013 г. № 412-ФЗ «Об аккредитации в национальной системе аккредитации».
98. Шаевич А.Б., Мучник В.Л., Кузьмин Н.М. // Ж. аналит. химии. 1990. Т. 45. С. 2251–2271.
99. ГОСТ Р 52361-2005. Контроль объекта аналитический. Термины и определения. ИПК Издательство стандартов. 2005.



100. Карнов Ю.А., Майоров И.А., Филимонов Л.Н. // Заводская лаборатория. 2001. Т. 67. № 2. С. 52–64.
101. Mitra S. (ed.). Sample Preparation Techniques in Analytical Chemistry. John Wiley & Sons. 2003. 458 p.
102. Карнов Ю.А., Савостин А.П. Методы пробоотбора и пробоподготовки. — М.: БИНОМ. Лаборатория знаний, 2009. — 243 с.
103. Барышников И.Ф., Попова Н.Н., Орбинская В.А. и др. Пробоотбирание и анализ благородных металлов. Справочник / под ред. И.Ф. Барышникова, 2-е изд. — М.: Металлургия, 1978. — 432 с.
104. Хан Г.А. Опробование и контроль технологических процессов обогащения. — М.: Недра, 1979. — 253 с.
105. Методические основы исследования химического состава горных пород, руд и минералов / ред. Г.В. Остроумов. — М.: Недра, 1979.
106. Анализ металлов. Пробоотбор. Справочник / ред. Ф. Энслин [и др.], пер. с нем. — М.: Металлургия, 1981. — 328 с.
107. Разумов В.А. Массовый анализ кормов. Справочник. — М.: Колос, 1982. — 176 с.
108. Каплан Б.Я., Филимонов Л.Н., Майоров И.А. Метрология аналитического контроля производства в цветной металлургии. — М.: Металлургия, 1989.
109. Gy P.M. Sampling of Particulate Materials. Theory and Practice. Elsevier, Amsterdam. 1979. 431 p.
110. Baule B., Benedette-Pichler A.A. // Z. Anal. Chem. 1928. V. 74. P. 442–456.
111. Gy P.M. Sampling of Heterogeneous and Dynamic Material Systems: Theories of Heterogeneity, Sampling and Homogenizing. Data Handling in Science and Technology. V. 10. Amsterdam: Elsevier, 1992. ISBN 0-444-89601-5.
112. Gy P.M. // Chemometrics and Intelligent Laboratory Systems. 2004. V. 74. P. 61–70.
113. Muskens P.J.W.M. // Anal.Chem.Acta. 1978. V. 103. P. 445–447.
114. Shewhart W.A. Statistical method from the view point of quality control. Pennsylvania: Lancaster Press, 1949. 197 p.
115. Murphy R.B. // Materials research and standards. 1961. V. 4. N. 1. P. 264–267.
116. Wernimont G.T. Validation of measurement process / Ed. J.R. De Voe. Wash.: Amer. Chem. Soc. 1977. P. 1–29 (ACS Symposium Ser.: N 63).
117. РМГ 76-2004 ГСИ. Внутренний контроль качества результатов количественного химического анализа. Стандартиформ, 2006.
118. Прохоров Ю.В., Розанов Ю.А. Теория вероятностей. — М.: Наука, 1987. — С. 352.
119. Shewhart W.A. Statistical method for the view point of quality control / Ed. Deming W.E. Pennsylvania: Lancaster Press, 1939. 155 p.
120. ГОСТ Р ИСО 7870-1-2011. Контрольные карты. — М.: Стандартиформ, 2012.
121. Levey S., Jennings E.R. // Am. J. Clin. Pathol. 1950. V. 20. P. 1059–1066.
122. Westgard J.O., Groth T., Aronsson T. et al. // Clin. Chem. 1977. V. 23. P. 1857–1867.
123. Westgard J.O., Groth T. // Clin. Chem. 1981. V. 27. P. 1536–1545.
124. Westgard J.O., Barry P.L., Hant M.R., Groth T.A. // Clin. Chem. 1981. V. 27. P. 493–501.
125. Westgard J.O., Westgard S.A. // Am.J.Clin.Path. 2006. V.125. P. 343–354.
126. Westgard J.O. // Ann.Clin.Biochem. 2003. V. 40. P. 593–611.
127. Мердок Дж. Контрольные карты. — М.: Финансы и статистика, 1986. — 150 с.
128. Hunter J.S. // J. of Qual. Technol. 1986. V. 18. P. 203–210.
129. Roberts S.W. // Technometrics. 1959. V. 1. N 3. P. 239–242.

130. *Hotteling H.* Multivariate quality control, illustrated by the air testing of sample bombsites. In: *Selected Techniques of Statistical Analysis*. Eisenhart C. et al., editors. New-York: McGraw-Hill, 1947. P. 111–184.
131. *Jackson J.E.* A User's Guide to Principal Components. NY: Wiley and Sons, 1991. 569 p.
132. *Miller P., Swanson R.* // *Appl. Math. And Comp. Sci.* 1998. V. 8. N 4. P. 775–792.
133. *Phatak A., ed.* // *NAmICS Newsletters*. 1999. N 19. P. 2–8.
134. *Клячкин В.Н.* Многомерный статистический контроль технологического процесса. — М.: Финансы и статистика, 2003. — 192 с.
135. *Aparisi F., Jabaloyes J., Carrion A.* // *Statistical properties of the |S| multivariate control chart* // *Comm. in Statistics – Theory and Methods*. 1999. V. 28. No. 11.
136. *ISO/IEC 17043:2010.* Conformity assessment — General requirements for proficiency testing.
137. *Dvorkin V.* // *Accreditation and quality assessment*. 1999. V. 4. N. 5. P. 207–211.
138. *Дворкин В.И.* // *Заводская лаборатория. Диагностика материалов*. 2008. Т. 74, № 12, С. 58–63.
139. *Gibbon G.A.* // *Laboratory Automation and Information Management issue*. V. 32/1. 1996. P. 1–6.
140. *Нуцков В.Ю.* // *Промышленные АСУ и контроллеры*. 2003. № 4. С. 48–50.
141. *Куцевич И.В.* / *Лабораторные информационные системы LIMS. Сборник статей: ООО «Маркетинг. Информационные технологии»*. 2006. С. 266–276.
142. *Нуцков В.Ю., Дюмаева И.В.* // *Заводская лаборатория*. 2004. № 10. С. 55–60.
143. *Савельев Е.В.* // *Партнеры и конкуренты*. 2005. № 4. С. 41–43.
144. *Меркуленко Н.Н.* / *Лабораторные информационные системы LIMS. Сборник статей: ООО «Маркетинг. Информационные технологии»*. 2006. С. 215–219.
145. *Самсонов А.В.* // *Промышленные АСУ и контроллеры*. 2004. № 10. С. 33–35.
146. *Терещенко А.Г., Терещенко В.А., Толстихина Т.В., Янин А.М.* // *Партнеры и конкуренты*. 2005. № 4. С. 44–45.
147. *Черепица С.В., Бычков С.М., Коваленко А.Н., Мазаник А.Л.* // *Партнеры и конкуренты*. 2005. № 1. С. 24–30.
148. *Каламбет Ю.А.* // *Пищевая промышленность*. 2005. № 3. С. 2–3.
149. *Дворкин В.И.* // *Партнеры и конкуренты*. 2000. № 4. С. 30–39.
150. *Р 50.2.003-2002.* Рекомендации по метрологии. ГСОЕИ. Внутренний контроль качества результатов измерений. Пакет программ QControl . ИПК Издательство стандартов, 2000 (с поправкой, опубл. в ИУС № 2, 2003 г.).
151. *Дворкин В.И.* // *Партнеры и конкуренты*. 2003. № 1. С. 26–39.
152. [www.qcontrol.ru](http://www.qcontrol.ru)