

Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Оренбургский государственный университет»

ОСНОВЫ ВЫЧИСЛИТЕЛЬНЫХ МЕТОДОВ В ХИМИИ

Учебное пособие

Рекомендовано ученым советом федерального государственного бюджетного образовательного учреждения высшего образования «Оренбургский государственный университет» для обучающихся по образовательным программам высшего образования по специальности 04.05.01 Фундаментальная и прикладная химия и по направлению подготовки 04.03.01 Химия.

Оренбург
2019

УДК 54:519.6(075.8)
ББК 24.1я73+22.19я73
О 75

Рецензент - кандидат физико-математических наук, доцент
А.Г. Четверикова

Авторы: О.Н. Каныгина, Е.В. Сальникова, П.А. Пономарева, Е.А. Осипова

О 75 Основы вычислительных методов в химии: учебное пособие / О.Н. Каныгина, Е.В. Сальникова, П.А. Пономарева, Е.А. Осипова; Оренбургский гос. ун-т. – Оренбург: ОГУ, 2019. –140 с.
ISBN

В пособии изложены математические основы обработки экспериментальных данных в приложении к задачам химической практики. Оно включает теоретический материал, примеры решения задач, вопросы для самоконтроля. Целью данного издания является активизация самостоятельной работы студентов.

Последовательность изложения и объем пособия соответствуют требованиям программы подготовки студентов по специальности 04.05.01 Фундаментальная и прикладная химия и направлению подготовки 04.03.01 Химия.

Данное пособие может быть полезно и для других нехимических специальностей при выполнении химического эксперимента и лабораторных работ.

Пособие будет полезно при выполнении курсовых и дипломных работ. Данным пособием также могут воспользоваться в научной работе студенты, аспиранты и преподаватели, специализирующиеся в области анализа объектов окружающей среды.

Коллектив авторов посвящает учебное пособие Стряпкову А.В.

УДК 54:519.6(075.8)
ББК 24.1я73+22.19я73

ISBN

© Каныгина О.Н.,
Сальникова Е.В.,
Пономарева П.А.,
Осипова Е.А., 2019
© ОГУ, 2019

Содержание

Введение	5
1 Погрешности химического анализа и их классификации	7
1.1 Классификация погрешностей по характеру вызывающих их причин	9
1.2 Классификация погрешностей по способу вычисления	13
1.3 Систематические погрешности косвенных измерений	15
1.4 Контрольные вопросы.....	28
2 Оценка случайных погрешностей с помощью методов математической статистики	30
2.1 Основные понятия классической статистики.....	30
2.2 Применение функции нормированного распределения Гаусса-Лапласа для обработки результатов химического анализа.....	37
2.3 Статистика малых выборок. Применение распределения Стьюдента для обработки результатов химического анализа.....	39
2.4 Проверка значений измерений на наличие грубых погрешностей (промахов)	48
2.5 Статистические методы как средства проверки моделей	52
2.6 Вычисление суммарной случайной погрешности результата химического анализа	59
2.7 Статистическая оценка предела обнаружения	64
2.8 Контрольные вопросы	67
3 Определение содержания компонента по градуировочному графику	69
3.1 Подбор параметров a и b методом наименьших квадратов для построения градуировочного графика	70
3.2 Определение содержания компонента с помощью уравнения линейной регрессии и градуировочного графика.....	79
3.3 Проверка линейности градуировочного графика	86
3.4 Контрольные вопросы	88

4 Применение метода наименьших квадратов для обработки нелинейных зависимостей	90
4.1 Обработка функций квадратичной зависимости	90
4.2 Обработка функций показательной зависимости	96
4.3 Обработка функций степенной зависимости	112
4.4 Контрольные вопросы.....	127
Список использованных источников	129
Приложение А (справочное) Функции Лапласа	131
Приложение Б (справочное) Значения коэффициента Стьюдента	132
Приложение В (справочное) Метод Крамера.....	134
Приложение Г (справочное) Значения Q-критерия	139
Приложение Д (справочное) Таблица значений критерия Фишера (F-критерия)..	140

Введение

Химическая наука и практика оперирует огромным количеством экспериментальных данных, результатов измерения физических величин химических веществ. Достоверность экспериментальных результатов оценивают, используя положения метрологии – науки об измерениях, методах и средствах обеспечения их единства и точности. Метрологический анализ правильности, воспроизводимости, предела обнаружения измеряемых величин выполняют с помощью методов математической статистики.

Математическая статистика – наука, которая опирается на теорию вероятностей и позволяет объективно оценивать результаты опытов и закономерности исследуемых явлений. В аналитической химии математическая статистика применяется, в основном, в двух аспектах: во-первых, для свертывания цифровой информации; во-вторых, для анализа экспериментального материала приемами, основанными на теории вероятностей. Оба указанных аспекта входят в круг вопросов, которые решают студенты естественных и технических факультетов, выпускники, магистранты, аспиранты, инженеры и методисты аналитических лабораторий.

Оценить достоверность полученных измерений и учесть погрешности эксперимента, можно пользуясь правилами и законами, изложенными в предлагаемом пособии, посвященном математической обработке результатов химического эксперимента. Это пособие, как показала апробация на практических занятиях, весьма полезно для студентов и после необходимой редакции и обновления материала должно оказывать существенную помощь в изучении практической химии. При изложении материала использована терминология и система обозначений, рекомендованная в нормативных документах.

В каждом из четырех разделов учебного пособия рассмотрены основные вопросы теории и показано ее практическое применение при решении типичных задач. К решениям задач даются краткие пояснения, показаны примеры

правильного представления результатов вычислений. В конце каждого раздела есть перечень вопросов для самоконтроля и задачи для самостоятельного решения.

Необходимый справочный материал в виде таблиц приведен в приложениях.

Учебное пособие предназначено для студентов химических и химико-технологических специальностей, но может быть полезным химикам-экспериментаторам, аспирантам и преподавателям, работающим в области анализа и исследования объектов окружающей среды.

1 Погрешности химического анализа и их классификации

Измеряемая величина – характерный признак явления, тела или вещества, который может выделяться качественно и определяться количественно. Значение физической величины – выражение размера физической величины в виде некоторого числа принятых для нее единиц, абсолютных или относительных. В аналитической химии такими единицами являются единицы массы (килограмм, грамм и так далее), проценты, г/дм³ и так далее.

Измерение величины – совокупность операций по применению технического средства, обеспечивающих нахождение соотношения (в явном или неявном виде) измеряемой величины с её единицей и получение значения этой величины. Таким образом, значение величины – это результат ее измерения. Измерения делятся на прямые и косвенные.

Прямое измерение – это измерение, при котором искомое значение физической величины получают непосредственно, например, измерение длины с помощью линейки, температуры с помощью термометра, времени секундомером и так далее. Косвенное измерение – это определение искомого значения физической величины на основании результатов измерения других физических величин, функционально связанных с искомой величиной. Исходя из этого определения, результат анализа относится к косвенным измерениям, определению содержания компонента в веществе на основе результатов измерения других физических величин, связанных с искомой через градуировочную функцию. Само измерение носит не прямой, а косвенный характер, поскольку измеряется не число частиц того или иного вида, а их совокупная масса – в гравиметрии, расход реагента – в титриметрии или другие физические параметры, пропорциональные массе. Обычный химический анализ многостадийен и включает ряд предварительных операций, предшествующих конечному определению.

Химический анализ включает ряд последовательных операций, в результате которых получают достоверные данные относительно качественного и

количественного состава исследуемого образца. К основным этапам химического анализа относят:

- 1) отбор представительной средней пробы;
- 2) переводение пробы в форму, удобную для анализа;
- 3) разделение компонентов пробы;
- 4) обнаружение компонента (качественный анализ) или определение его количества;
- 5) вычисления и обработку результатов анализа.

При **обнаружении** компонента (этап 4) фиксируют лишь появление аналитического сигнала: выпадение осадка, изменение окраски, появление линии в спектре и так далее.

При определении количества компонента **измеряется величина аналитического сигнала**: объем титранта, масса осадка, оптическая плотность, сила тока, интенсивность линии в спектре и так далее, используя как простые измерительные средства (бюретки, аналитические весы), так и сложное техническое оборудование (фотоэлектроколориметр, полярограф, спектрограф, потенциометр). После этого рассчитывают содержание компонента (этап 5), используя функциональную зависимость «аналитический сигнал – содержание компонента», которая устанавливается заранее расчетным или опытным путем и может быть представлена в виде формулы, таблицы или графика. Например, в прямых титриметрических методах анализа содержание компонента X находят по формуле:

$$m(X) = \frac{c(f_{\text{экв.}}(A)A) \cdot V(A_p) \cdot M(f_{\text{экв.}}(X)X)}{1000}, \quad (1)$$

где $m(X)$ – масса определяемого компонента, г.

Содержание определяемого компонента может быть выражено в единицах массы (г или мг), количестве вещества (моль), или концентрации.

На каждом этапе химического анализа измерения производятся с определенными погрешностями и конечный результат анализа также имеет свою погрешность. Конечной стадией количественного анализа химического состава исследуемого образца является обработка результатов эксперимента (не менее сложная задача, чем само проведение измерений), которая включает изучение литературы, дополнительные исследования для выяснения причин возникновения погрешностей и статистическую обработку полученных данных.

Рассмотрим типы погрешностей, возникающих при проведении химического анализа. Их можно классифицировать по характеру причин, которые их вызывают, по способу вычисления и так далее.

1.1 Классификация погрешностей по характеру вызывающих их причин

К ним относятся: систематические, случайные и грубые (промахи).

К *систематическим* относят погрешности, постоянные по знаку и по величине для всех измерений. Они обусловлены наличием постоянно действующих причин: инструментальных, методических, реактивных и т. д. Систематические инструментальные возникают при неправильной градуировке приборов (мерных колб, бюреток, пипеток, весов, фотоколориметров, рН-метров) и встречаются в основном, на четвертом этапе химического анализа. Периодическая поверка приборов сводит инструментальные систематические погрешности к минимуму.

Основной вклад в систематическую погрешность вносят методические погрешности, возникающие при отборе пробы, переведении пробы в удобную для анализа форму (растворение, сплавление, спекание) и погрешности, возникающие при разделении компонентов. Так, в гравиметрическом анализе погрешности обусловлены частичной растворимостью осаждаемой формы, процессами соосаждения, отклонением от стехиометрического состава гравиметрической формы. В титриметрическом анализе типичной методической погрешностью

является индикаторная ошибка кислотно-основного титрования из-за несовпадения pT индикатора со значением pH в точке эквивалентности; капельная ошибка титрования.

Хорошие методики химического анализа не должны содержать существенных методических погрешностей, или, в крайнем случае, специально указывать на наиболее слабые (с точки зрения методики) моменты анализа, с тем, чтобы аналитик их выполнял особенно тщательно.

Систематическая реактивная погрешность обусловлена примесями, содержащимися в реактивах, воде и других растворителях. Наиболее простой и распространенный способ нахождения величины реактивной погрешности состоит в параллельном определении содержания компонента в исследуемой и «холостой» пробах. «Холостая» проба должна содержать все те же компоненты, что и исследуемая, за исключением определяемого компонента. Вычитание результата анализа «холостой» пробы из результата анализа исследуемого образца позволяет исключить реактивную погрешность.

Правильно поставленная методика устраняет систематическую реактивную погрешность. Например, при фотометрическом определении используют «холостой» раствор как раствор сравнения для устранения систематической реактивной погрешности. Можно автоматически устранить систематическую погрешность путем релятивизации (англ. *relative* – относительный). Например, если отбор стандартного и исследуемого растворов проводить с помощью одной и той же пипетки, систематическая погрешность из-за неправильной калибровки пипетки устраняется.

Систематические погрешности приводят к неправильным результатам. Под правильностью анализа понимают близость полученного результата к истинному значению.

Общепринятый прием оценки правильности метода – анализ стандартного образца (образца с точно известным содержанием определяемого компонента). Обязательное условие применения стандартного образца в химическом анализе – максимальная близость состава и свойств стандартного образца и анализируемой

пробы. Для оценки правильности метода проводят многократный химический анализ образца и сравнивают найденное содержание с паспортным (действительным) содержанием определяемого компонента. В титриметрических методах анализа широко используют фиксаналы, стандартные образцы для приготовления стандартных растворов.

Таким образом, выявление систематических погрешностей требует специальных исследований. Выявленные систематические погрешности можно устранить или оценить их величину и внести соответствующую поправку в результат исследования.

Случайные погрешности в отличие от систематических не имеют видимой причины (небольшое изменение температуры, влажности воздуха, давления, случайные потери вещества, случайные загрязнения посуды и так далее). Каждая из них незначительно влияет на общий результат анализа, поэтому индивидуальное рассмотрение не имеет смысла и оценивается общая случайная погрешность. Большие случайные погрешности менее вероятны, чем маленькие.

Основным признаком наличия случайных погрешностей является несовпадение результатов повторных анализов. Степень близости результатов единичных определений друг к другу (рассеяние единичных результатов относительно среднего значения) характеризует воспроизводимость анализа.

Случайные погрешности нельзя устранить, но можно уменьшить при многократном повторении опыта. Общую случайную погрешность оценивают путем статистической обработки параллельных результатов. Статистически можно обрабатывать как параллельно измеренные величины (значения массы, объема, показаний прибора), так и конечные результаты анализа.

Основы статистической обработки результатов химического анализа. Сначала выявляют систематические погрешности. Например, систематическая погрешность, обусловленная неправильными показаниями разных приборов, при измерении аналитического сигнала в разных лабораториях переходит в случайную. Прием, который позволяет перевести систематические погрешности в

случайные, называется рандомизацией систематических погрешностей (англ. *random* –случайно).

Промах – это погрешность, обусловленная неисправностью измерительных приборов, неправильным отсчетом показаний, резким изменением условий при выполнении анализа, грубыми ошибками аналитика и т.п., резко искажающая результат анализа.

Схема, поясняющая понятия систематических и случайных погрешностей, представлена на рисунке 1.

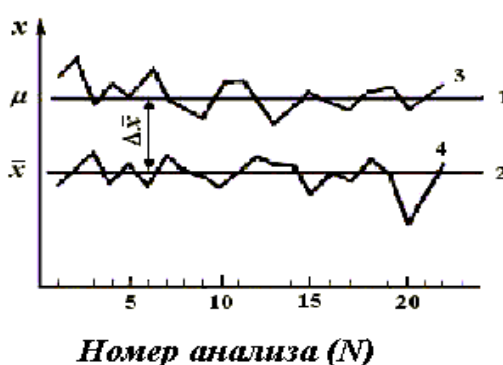


Рисунок 1 - Систематические и случайные погрешности при многократном химическом анализе

Прямая 1 отвечает идеальному случаю, когда во всех определениях (N) отсутствуют систематические и случайные погрешности. Линии 2 и 3 – также идеализированные примеры химического анализа. В случае 2 полностью отсутствуют случайные погрешности, но все N определений имеют отрицательную систематическую погрешность $\Delta \bar{x}$; в случае 3 полностью отсутствует систематическая погрешность. Реальную ситуацию отображает линия 4: имеются как случайные, так и систематические погрешности.

1.2 Классификация погрешностей по способу вычисления

По способу вычисления погрешности классифицируют на абсолютные и относительные.

Абсолютная погрешность единичного определения равняется разности между полученным результатом (x_i) и **истинным** значением (μ) определяемой величины:

$$\Delta x_i = x_i - \mu \quad (2)$$

Поскольку истинное значение неизвестно, то полученный результат x_i сравнивают с **действительным** результатом, который принимают за истинный. Действительное значение – это экспериментальное или рассчитанное значение, которое настолько близко к истинному, что может быть использовано вместо него. Например, содержание определяемого компонента в стандартном образце. Действительный результат, в свою очередь, может быть недостоверным, поэтому оценить истинную абсолютную погрешность довольно трудно.

Пример 1.1 Взвешивание образца, действительная масса которого равна 0,1000 г, показало, что его масса равняется 0,1001 г. Определить абсолютную погрешность взвешивания.

Решение

Абсолютную погрешность взвешивания определяем по формуле:

$$\Delta x_i = x_i - \mu .$$

$$\Delta x_i = 0,1001 \text{ г} - 0,1000 \text{ г} = 0,0001 \text{ г}$$

Ответ: $\Delta x_i = 0,0001 \text{ г}$

Абсолютная погрешность выражается в тех же самых единицах, что и измеряемая величина (г, дм³ и т.п.).

Абсолютная погрешность среднего значения равняется разности между средним значением и истинным значением определяемой величины:

$$\Delta \bar{x} = \bar{x} - \mu$$

Пример 1.2 Действительное значение массовой доли (%) хлорид-ионов в образце равно 24,33. Среднее значение – 24,39. Определить абсолютную погрешность среднего значения массовой доли хлорид-ионов в образце.

Решение

Абсолютную погрешность среднего значения массовой доли (%) хлорид-ионов в образце определяем по формуле:

$$\Delta \bar{x} = \bar{x} - \mu; \quad \Delta \bar{x} = 24,39 - 24,33 = 0,06 .$$

Ответ: $\Delta \bar{x} = 0,06 \%$.

Относительная погрешность определения равна отношению абсолютной погрешности $(x_i - \mu)$ или $(\bar{x} - \mu)$ к истинному значению (μ) определяемой величины и, как правило, не имеет знака.

Относительная погрешность может быть выражена в долях, но чаще ее выражают в процентах:

$$\frac{|\Delta x|}{\mu} \cdot 100 \quad \text{или} \quad \frac{|\Delta \bar{x}|}{\mu} \cdot 100 .$$

Для рассмотренного выше примера 1.1 относительная погрешность (%) единичного определения равна:

$$\frac{|\Delta x_i|}{\mu} \cdot 100 = \frac{0,0001}{0,1000} \cdot 100 = 0,1 .$$

Для примера 1.2 относительная погрешность (%) определения среднего равна:

$$\frac{|\Delta \bar{x}|}{\mu} \cdot 100 = \frac{0,06}{24,36} \cdot 100 = 0,2 .$$

В зависимости от того, завышает или занижает погрешность результат анализа, погрешности могут быть положительными или отрицательными.

В примере 1.2 среднее значение массовой доли (%) хлорид-ионов составляет 24,19, а абсолютная погрешность будет отрицательной величиной:

$$\Delta \bar{x} = 24,19 - 24,33 = - 0,14 .$$

Знак «-» указывает на то, что результат анализа занижен в сравнении с действительным значением.

Погрешность можно оценить для единичного измерения, для среднего значения, найденного из нескольких параллельных определений, и для метода анализа в целом. Для оценки случайной погрешности единичного измерения используют величину стандартного отклонения для единичного (s) или доверительный интервал для единичного измерения.

Для оценки случайной погрешности среднего значения используют величину стандартного отклонения для среднего значения ($s_{\bar{x}}$) или доверительный интервал для среднего значения.

Для оценки случайной погрешности метода анализа в целом используют величину стандартного отклонения для единичного измерения (s) или дисперсию (s^2).

В литературе по химическому анализу встречается также термин «предельная» погрешность, сущность которой будет рассмотрена ниже.

1.3 Систематические погрешности косвенных измерений

Результаты большинства химико-аналитических определений получают путем косвенных измерений, например, даже при вычислении массовой доли

компонента в пробе в титриметрическом анализе проводят несколько операций умножения и деления:

$$w(\%)(X) = \frac{c(f_{\text{экв.}}(A)A) \cdot \bar{V}(A_{\text{расч.}}) \cdot M(f_{\text{экв.}}(X)X)}{1000} \times \frac{V_{\text{м.к.}}}{V^a(X_{\text{расч.}})} \cdot \frac{100}{m_{\text{нав.}}} \quad (3)$$

Очевидно, что погрешность определения функции $w(\%)(X)$ будет зависеть от погрешностей всех аргументов: $c(f_{\text{экв.}}(A)A)$; $\bar{V}(A_{\text{расч.}})$; $M(f_{\text{экв.}}(X)X)$; $V_{\text{м.к.}}$; $V^a(X_{\text{расч.}})$; $m_{\text{нав.}}$.

Пусть y – результат косвенного химико-аналитического определения; x_1, x_2, \dots, x_n – аргументы, значения которых измеряются в ходе химического анализа, f – определенная функциональная зависимость, которая связывает их:

$$y = f(x_1, x_2, \dots, x_n).$$

В общем виде задачу нахождения суммарной погрешности можно сформулировать таким образом. Если известны погрешности $(\pm \Delta x_i)$, допущенные при измерении отдельных аргументов, и вид функции f , можно оценить погрешность $(\pm \Delta y)$ в определении величины y .

Для полного решения задачи используется математический аппарат дифференциального исчисления. Способ вычисления суммарной предельной погрешности определяется видом погрешности (систематическая или случайная) и видом арифметических действий над экспериментальными значениями.

Различают правила вычисления суммарной предельной систематической погрешности и суммарной предельной случайной погрешности.

Ниже приведены четыре правила нахождения суммарной систематической погрешности.

Правило 1. Для функции вида $y = x_1 + x_2$ или $y = x_1 - x_2$, $|\Delta y| = |\Delta x_1| + |\Delta x_2|$.

Абсолютная погрешность суммы (или разности) равна сумме абсолютных погрешностей переменных.

Правило 2. Для функции вида $y = x_1 \cdot x_2$ или $y = \frac{x_1}{x_2}$,

$$\left| \frac{\Delta y}{y} \right| = \left| \frac{\Delta x_1}{x_1} \right| + \left| \frac{\Delta x_2}{x_2} \right|.$$

Относительная погрешность произведения или частного от деления равна сумме относительных погрешностей аргументов. Погрешность суммы или разности определяется абсолютными величинами погрешностей, а погрешность произведения и частного от деления – относительными величинами.

Правило 3. Для функции вида $y = x^n$

$$\left| \frac{\Delta y}{y} \right| = \left| n \cdot \frac{\Delta x}{x} \right|.$$

Относительная погрешность степени равняется относительной погрешности величины, которая возводится в степень, умноженной на показатель степени.

Аналогично, для функции вида $y = \sqrt[n]{x}$:

$$\left| \frac{\Delta y}{y} \right| = \left| \frac{1}{n} \cdot \frac{\Delta x}{x} \right|.$$

Правило 4. Для функции вида $y = \lg x$

$$|\Delta y| = 0,4343 \left| \frac{\Delta x}{x} \right|.$$

Абсолютная погрешность логарифма (по основанию 10) равна относительной погрешности величины, которая логарифмируется, умноженной на 0,4343.

При вычислении антилогарифма: $x = \text{antlg } y$

$$\left| \frac{\Delta x}{x} \right| = \left| \frac{\Delta y}{0,4343} \right|$$

Относительная погрешность антилогарифма равна абсолютной погрешности логарифма, разделенной на 0,4343.

Если предельные абсолютные систематические погрешности отдельных измерений неизвестны, то за значение $\pm \Delta x$ принимают значение минимального деления шкалы прибора, иногда половины или даже десятой доли деления шкалы прибора. Например, предельная абсолютная систематическая погрешность навески на аналитических весах принимается равной $\pm 0,0001$ г (значение одного деления

шкалы), на теххимических – $\pm 0,01$ г (или $\pm 0,005$ г). Абсолютная систематическая погрешность определения объема раствора по шкале бюретки объемом $25,00 \text{ см}^3$ составляет $\pm 0,05 \text{ см}^3$ (половина деления шкалы) (иногда используют даже значение $\pm 0,01 \text{ см}^3$), микробюретки объемом 1; 2; 3 см^3 составляет $\pm 0,005 \text{ см}^3$ (иногда используют даже значение $\pm 0,001 \text{ см}^3$).

Предельные систематические погрешности мерных колб и пипеток класса А (первого класса) приведены в таблице 1.

Рассмотренные правила справедливы для нахождения предельных систематических погрешностей, если известны максимально возможные погрешности отдельных измерений. Формулы суммарных предельных систематических погрешностей для функций, чаще всего встречающихся в аналитической химии, приведены во втором столбце таблицы 2. Вычисление суммарной погрешности с определенным знаком проводится по формулам, приведенным в третьем столбце таблицы 2.

Таблица 1 - Предельные систематические погрешности мерной посуды

Мерные колбы класса А					
Объем, см^3	2000	1000	500	250	100
Максимально допустимая абсолютная погрешность, см^3	0,6	0,4	0,25	0,15	0,10
Максимально допустимая относительная погрешность, %	0,03	0,04	0,05	0,06	0,10
Пипетки класса А					
Объем, см^3	100	50	25	10	2
Максимально допустимая абсолютная погрешность, см^3	0,080	0,050	0,030	0,020	0,005
Максимально допустимая относительная погрешность, %	0,08	0,10	0,012	0,20	0,25

Таблица 2 - Систематические погрешности косвенных измерений

Функция	Предельная суммарная погрешность	Суммарная погрешность с учетом знака
$y = x_1 + x_2$	$ \Delta y = \Delta x_1 + \Delta x_2 $	$\Delta y = \Delta x_1 + \Delta x_2$
$y = x_1 - x_2$	$ \Delta y = \Delta x_1 + \Delta x_2 $	$\Delta y = \Delta x_1 - \Delta x_2$
$y = x_1 \cdot x_2$	$\left \frac{\Delta y}{y} \right = \left \frac{\Delta x_1}{x_1} \right + \left \frac{\Delta x_2}{x_2} \right $	$\frac{\Delta y}{y} = \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2}$
$y = \frac{x_1}{x_2}$	$\left \frac{\Delta y}{y} \right = \left \frac{\Delta x_1}{x_1} \right + \left \frac{\Delta x_2}{x_2} \right $	$\frac{\Delta y}{y} = \frac{\Delta x_1}{x_1} - \frac{\Delta x_2}{x_2}$
$y = \frac{x_1 \cdot x_2}{x_3}$	$\left \frac{\Delta y}{y} \right = \left \frac{\Delta x_1}{x_1} \right + \left \frac{\Delta x_2}{x_2} \right + \left \frac{\Delta x_3}{x_3} \right $	$\frac{\Delta y}{y} = \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2} - \frac{\Delta x_3}{x_3}$
$y = x^n$	$\left \frac{\Delta y}{y} \right = \left n \cdot \frac{\Delta x}{x} \right $	$\frac{\Delta y}{y} = n \cdot \frac{\Delta x}{x}$
$y = \ln x$	$ \Delta y = \left \frac{\Delta x}{x} \right $	$\Delta y = \frac{\Delta x}{x}$
$y = \lg x$	$ \Delta y = 0,4343 \left \frac{\Delta x}{x} \right $	$\Delta y = 0,4343 \cdot \frac{\Delta x}{x}$
$x = \text{antlg } y$	$\left \frac{\Delta x}{x} \right = \left \frac{\Delta y}{0,4343} \right $	$\frac{\Delta x}{x} = \frac{\Delta y}{0,4343}$

Примеры нахождения суммарных систематических погрешностей.

Пример 1.3 Найдите предельную систематическую погрешность при приготовлении раствора сульфата меди с концентрацией $40,00 \text{ г/дм}^3$ растворением навески кристаллогидрата $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ массой $10,000 \text{ г}$ в колбе объемом 250 см^3 .

Решение

1. Какова абсолютная предельная погрешность измерения объема ?

В соответствии с данными таблицы 1 $\Delta V = \pm 0,15 \text{ см}^3$.

2. Какова предельная относительная погрешность измерения объема ?

$$\left| \frac{\Delta V}{V} \right| = \frac{0,15}{250,0} = 6,0 \cdot 10^{-4}.$$

3. Какова абсолютная предельная погрешность взятия навески ?

Судя по требуемой точности взвешивания навески (10,000 г), взвешивание производят на теххимических весах, в этом случае предельную погрешность взвешивания (Δm), как правило, принимают равной $\pm 0,005$ г.

4. Какова предельная относительная погрешность взвешивания ?

$$\left| \frac{\Delta m}{m} \right| = \frac{0,005}{10,000} = 5,0 \cdot 10^{-4}.$$

5. Какова концентрация (г/дм³) раствора сульфата меди ?

$$c(\text{CuSO}_4 \cdot 5\text{H}_2\text{O}) = \frac{m(\text{CuSO}_4 \cdot 5\text{H}_2\text{O})}{V(\text{CuSO}_4 \cdot 5\text{H}_2\text{O} \text{ расм.})} = \frac{10,000}{0,2500} = 40,00 \text{ г/дм}^3.$$

6. Какова предельная относительная погрешность приготовления раствора с заданной концентрацией ?

$$\left| \frac{\Delta c}{c} \right| = \left| \frac{\Delta m}{m} \right| + \left| \frac{\Delta V}{V} \right| = 5,0 \cdot 10^{-4} + 6,0 \cdot 10^{-4} = 11,0 \cdot 10^{-4}.$$

7. Какова предельная абсолютная погрешность приготовления раствора с заданной концентрацией ?

$$\Delta c = \left| \frac{\Delta c}{c} \right| \cdot c = 11,0 \cdot 10^{-4} \cdot 40,00 = 0,044 \approx 0,04 \text{ г/дм}^3;$$

$$\Delta c = \pm 0,04 \text{ г/дм}^3.$$

Значение предельной абсолютной погрешности при приготовлении раствора указывает на то, что недостоверность в значении концентрации раствора имеется уже во втором десятичном знаке после запятой. Поэтому значение концентрации раствора $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ надо записать с двумя десятичными знаками после запятой.

$$\text{Ответ: } c(\text{CuSO}_4 \cdot 5\text{H}_2\text{O}) = 40,00 \text{ г/дм}^3; \quad \Delta c = \pm 0,04 \text{ г/дм}^3.$$

Пример 1.4 Найдите абсолютную и относительную погрешности при определении общей массы нескольких изделий из платины. Масса каждого изделия (г) была определена взвешиванием с разной систематической погрешностью: платинового тигля 4,05 (+0,01), платиновой чашки 27,84 (+0,02), крышки тигля 2,18 (-0,03), наконечника к щипцам 3,44 (+0,01). (В скобках указаны абсолютные погрешности результатов взвешивания отдельных изделий с учетом знака).

Решение

1. Какова абсолютная погрешность определения общей массы ?

$$\Delta m = 0,01 + 0,02 + (-0,03) + 0,01 = 0,01 \text{ г}.$$

2. Какова общая масса изделий ?

$$m = 4,05 + 27,84 + 2,18 + 3,44 = 37,51 \text{ г}.$$

3. Какова относительная погрешность при определении общей массы ?

$$\left| \frac{\Delta m}{m} \right| = \frac{0,01}{37,51} = 2,7 \cdot 10^{-4} \text{ или } 2,7 \cdot 10^{-2} \%$$

Ответ: $\Delta m = 0,01 \text{ г}; \frac{\Delta m}{m} = 2,7 \cdot 10^{-2} \%$

Значащие цифры и правила округления. Измеряя любую величину, исследователь получает определенное число, считывая его со шкалы измерительного прибора. При этом количество цифр (разрядов) числа не может быть произвольным. Оно определяется размахом шкалы и ценой одного ее деления. На следующем этапе проводится вычисление результатов анализа с использованием формул, аргументами в которых являются измеренные величины. Результаты анализа и вычислений принято выражать только значащими цифрами, которыми называют все достоверные цифры плюс первую из недостоверных (сомнительных). Все результаты надо округлять до первой недостоверной цифры.

Определить, какие цифры являются достоверными не всегда просто. Для этого надо знать величину случайной и систематической погрешности каждого результата, то есть его недостоверность.

В качестве статистических критериев оценки случайной погрешности для серии параллельных измерений можно использовать стандартное отклонение и доверительный интервал.

Недостоверность единичных измерений нельзя оценить с помощью статистических критериев. Для единичных измерений недостоверность оценивают величиной максимально возможной систематической погрешности, за значение которой принимают цену наименьшего деления шкалы прибора.

Пример 1.5 На титрование навески технического препарата буры ($\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$) массой 0,5866 г израсходован раствор соляной (хлороводородной) кислоты объемом 15,5 см³ с молярной концентрацией вещества HCl 0,2000 моль/дм³. Сколько значащих цифр необходимо оставить в результате вычисления массовой доли $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ в анализируемом препарате ?

Решение

1. Какова массовая доля $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ в исследуемом образце ?

$$w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = \frac{c(\text{HCl}) \cdot V(\text{HCl}) \cdot M(1/2\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})}{1000} \cdot \frac{100}{m_{\text{нав.}}} ;$$

$$M(1/2\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = 190,70 \text{ г/моль} ;$$

$$w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = \frac{0,2000 \cdot 15,5 \cdot 190,70}{1000} \cdot \frac{100}{0,5866} = 98,503 .$$

2. Какова предельная абсолютная недостоверность определения каждой из измеренных величин?

$$\Delta V(\text{HCl}) = \pm 0,05 \text{ дм}^3 \text{ (половина цены деления бюретки)};$$

$$\Delta m(\text{нав.}) = \pm 0,0001 \text{ г (цена деления шкалы вейтографа весов)} ;$$

$$\Delta c(\text{HCl}) = \pm 0,0001 \text{ моль/дм}^3 ;$$

$$\Delta M(1/2 \text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = \pm 0,01 \text{ г/моль}^1).$$

3. Какова предельная относительная недостоверность (%) определения каждой из измеренных величин?

$$\frac{\Delta V(\text{HCl})}{V(\text{HCl})} \cdot 100 = \frac{0,05 \cdot 100}{15,5} = 0,33 ;$$

$$\frac{\Delta m(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})}{m_{\text{нав.}}} \cdot 100 = \frac{0,0001 \cdot 100}{0,5866} = 0,017 ;$$

$$\frac{\Delta c(\text{HCl})}{c(\text{HCl})} \cdot 100 = \frac{0,0001 \cdot 100}{0,2000} = 0,050;$$

¹⁾ Поскольку сведения о недостоверности значений молярной концентрации и молярной массы эквивалента вещества $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ в условии задачи отсутствуют, *принимают недостоверность их определения равной ± 1 в последней известной значащей цифре.*

$$\frac{\Delta M(1/2 \text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})}{M(1/2 \text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})} \cdot 100 = \frac{0,01 \cdot 100}{190,70} = 0,0052 .$$

4. Каково суммарное значение предельной относительной недостоверности (%) результата анализа ?

$$\begin{aligned} \frac{\Delta w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})}{w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})} &= \\ &= \frac{\Delta V(\text{HCl})}{V(\text{HCl})} \cdot 100 + \frac{\Delta m(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})}{m_{\text{нав.}}} \cdot 100 + \frac{\Delta c(\text{HCl})}{c(\text{HCl})} \cdot 100 + \frac{\Delta M(1/2 \text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})}{M(1/2 \text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})} . \\ \frac{\Delta w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})}{w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O})} &= 0,33 + 0,017 + 0,05 + 0,0052 = 0,402 . \end{aligned}$$

5. Каково суммарное значение предельной абсолютной недостоверности результата анализа ?

$$\Delta w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = \frac{0,402 \cdot 98,503}{100} = 0,396^{2)} \approx 0,4 ;$$

$$\Delta w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = \pm 0,4 .$$

6. Сколько значащих цифр необходимо оставить в результате определения массовой доли $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ в образце ?

Значение $\Delta w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = 0,4$ показывает, что недостоверной цифрой в результате анализа ($w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = 98,503$) будет уже первая цифра после запятой в десятичной дроби, поэтому результат анализа необходимо округлить до 98,5.

$$\text{Ответ: } w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = 98,5 ; \quad \Delta w(\%)(\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}) = \pm 0,4 .$$

В данном примере рассмотрен строгий подход для оценки значности (числа значащих цифр) результата, основанный на вычислении суммарной недостоверности результата анализа. На практике аналогичные вычисления, как правило, не проводят, а используют правило 1:

²⁾ **Правило дополнения** -если первая из цифр, которая отбрасывается, есть 5, 6, 7, 8 или 9, то предыдущую цифру при округлении надо увеличить на 1. **Правило четности**: при округлении цифры 5, если предпоследняя цифра **четная**, то цифра 5 отбрасывается, если она **нечетная**, то предпоследняя цифра увеличивается на 1.

Правило 1. *Значность произведения или частного от деления определяется значностью сомножителя с наименьшим числом значащих цифр.*

Если пользоваться этим нестрогим правилом, то в ответе надо оставить четыре значащих цифры и результат записать как 98,50, поскольку в числах 0,2000 и 15,15 содержится по четыре значащих цифры, а в числе 190,70 – пять. Хотя при использовании этого правила есть риск ошибиться, как в примере 1.5, все-таки в большинстве случаев пользуются именно этим правилом.

Примечание - Нуль в числах может быть значащим и незначащим. Нули, стоящие в десятичной дроби в начале числа, всегда незначащие и служат лишь для указания места запятой в десятичной дроби. Например, число 0,02 содержит одну значащую цифру (2). Нули, стоящие между цифрами, всегда значащие. Например, в числе 0,301 три значащих цифры.

Нули в конце числа могут быть значащими и незначащими. Нули, стоящие в конце числа после запятой в десятичной дроби, считаются значащими. Например, в числе 190,70 пять значащих цифр. Нули, стоящие в конце целого числа, могут быть значащими, а могут просто указывать на порядок числа. Например, в числе 300 может быть одна значащая цифра (цифра 3), две значащих цифры (цифры 3 и 0) и три значащих цифры (цифры 3, 0 и 0). Во избежание неопределенности, рекомендуется в таких случаях представить число в виде произведения, которое содержит только значащие цифры. Например, если число 300 можно записать как $3 \cdot 10^2$ (одна значащая цифра), если две значащих цифры – $3,0 \cdot 10^2$, если три значащих цифры – $3,00 \cdot 10^2$.

Выше было рассмотрено нестрогое правило для оценки значности произведения или частного от деления. Рассмотрим правила 2-6 определения значности при других арифметических действиях.

Правило 2. *При сложении или вычитании значность суммы или разности определяется значностью числа, имеющего наименьшее количество десятичных знаков (имеющего наибольшую абсолютную не Достоверность). Числа, имеющие степени, преобразуют, приводя показатели степеней слагаемых к наибольшему.*

Пример 1.6 Сколько значащих цифр необходимо оставить в результате сложения таких значений концентраций гидроксид-ионов (моль/дм³): $4,00 \cdot 10^{-2}$; $5,55 \cdot 10^{-3}$ и $1 \cdot 10^{-6}$.

Решение

1. Преобразуем числа, приводя показатели степеней к наибольшему. Получим:

$$4,00 \cdot 10^{-2}; \quad 0,555 \cdot 10^{-2}; \quad 0,0001 \cdot 10^{-2}.$$

2. Какова сумма всех значений концентраций гидроксид-ионов (моль/дм³)?

$$4,00 \cdot 10^{-2} + 0,555 \cdot 10^{-2} + 0,0001 \cdot 10^{-2} = 4,5551 \cdot 10^{-2}.$$

Учтем, что число $4,00 \cdot 10^{-2}$ имеет наименьшее количество десятичных знаков (наибольшую абсолютную недостоверность) – уже во втором десятичном знаке после запятой, поэтому результат (сумму) необходимо выразить числом, которое содержит две цифры после запятой.

$$c(\text{OH}^-) = 4,56 \cdot 10^{-2} \text{ моль/дм}^3.$$

Ответ: $c(\text{OH}^-) = 4,56 \cdot 10^{-2} \text{ моль/дм}^3$.

Правило 3. При возведении числа в степень относительная недостоверность увеличивается в число раз, равное показателю степени.

Так при возведении числа в квадрат относительная недостоверность увеличивается в два раза. При извлечении корня также необходимо учитывать относительную недостоверность подкоренного числа.

Правило 4. Относительная недостоверность результата извлечения квадратного корня вдвое меньше относительной недостоверности подкоренного числа.

Пример 1.7 Извлечь квадратный корень из числа 10,00 и определить число значащих цифр в результате.

Решение

1. Какова абсолютная недостоверность числа 10,00, стоящего под корнем ?

При определении абсолютной недостоверности числа, если нет специальных указаний, допускают, что последняя цифра может изменяться в интервале ± 1 . Абсолютная недостоверность числа 10,00 равняется 0,01.

2. Какова относительная недостоверность числа 10,00, стоящего под корнем ?

$$\frac{0,01}{10,00} = 0,001.$$

3. Какова относительная недостоверность результата извлечения корня ?

$$0,001 : 2 = 0,0005 .$$

4. Каково значение квадратного корня из числа 10,00 ?

$$\sqrt{10,00} = 3,1623 .$$

5. Какова абсолютная недостоверность результата извлечения корня ?

$$0,0005 \cdot 3,1623 \cong 0,002 .$$

6. Сколько значащих цифр имеет результат извлечения квадратного корня ?

Поскольку абсолютная недостоверность результата извлечения корня в третьем десятичном знаке после запятой, полученный результат округляем до третьего десятичного знака после запятой.

$$\sqrt{10,00} = 3,162 .$$

Ответ: $\sqrt{10,00} = 3,162$; в результате необходимо указывать четыре значащих цифры.

Правило 5. При логарифмировании числа количество значащих цифр в мантиссе равняется количеству цифр, которое содержал нестепенной член числа. Характеристика логарифма не входит в количество значащих цифр, поскольку она указывает лишь на порядок числа, которое логарифмируют.

Пример 1.8 Вычислить pH раствора, если молярная концентрация ионов водорода в нем равняется $6,6 \cdot 10^{-11}$ моль/дм³. Определить количество значащих цифр в результате.

Решение

1. Каково значение pH раствора ?

$$\text{pH} = -\lg c(\text{H}^+) = -\lg 6,6 \cdot 10^{-11} = 11 - 0,8195 = 10,1805 .$$

2. Сколько значащих цифр в результате ?

Первые две цифры (1 и 0) в результате получаются при логарифмировании степенного члена 10^{-11} и определяют характеристику логарифма. Нестепенной член числа, которое логарифмировали, содержал две значащих цифры (6 и 6), поэтому мантисса должна содержать также две значащих цифры (1 и 8).

Ответ: pH = 10,18. Количество значащих цифр выросло с двух до четырех.

Правило 6. При вычислении антилогарифма количество значащих цифр в нестепенном члене полученного числа равно количеству цифр в мантиссе.

Пример 1.9 pH раствора равен 5,63. Вычислить молярную концентрацию ионов водорода в растворе. Определить количество значащих цифр в результате.

Решение

1. Какова концентрация ионов водорода в растворе ?

$$\text{pH} = -\lg c(\text{H}^+) ; \lg c(\text{H}^+) = -\text{pH} .$$

$$c(\text{H}^+) = \text{antlg} (-\text{pH}); \quad c(\text{H}^+) = \text{antlg} (-5,63) = 2,344 \cdot 10^{-6} .$$

2. Сколько значащих цифр в результате ?

Мантисса логарифма содержала две значащих цифры (6 и 3), поэтому нестепенной член результата также должен содержать две значащих цифры (2 и 3). Поэтому число $2,344 \cdot 10^{-6}$ округляем до $2,3 \cdot 10^{-6}$.

Ответ: $c(\text{H}^+) = 2,3 \cdot 10^{-6}$ моль/дм³. Количество значащих цифр при вычислении антилогарифма уменьшилось с трех до двух.

Из примеров 1.8 и 1.9 видно, что при вычислении логарифмов количество значащих цифр возрастает, а при вычислении антилогарифмов количество значащих цифр уменьшается.

1.4 Контрольные вопросы

1. Привести примеры аналитических сигналов, измерение которых лежит в основе методов химического анализа.
2. Привести классификацию погрешностей по характеру причин, которые их вызывают.
3. Дать определения систематической и случайной погрешностям.
4. Перечислить источники систематических погрешностей в гравиметрическом и титриметрическом методах анализа.
5. Привести примеры оценки систематических погрешностей бюретки, пипетки, мерной колбы. Как уменьшить систематическую погрешность пипетки?
6. Указать причины появления систематических и случайных погрешностей взвешивания:
 - а) поглощение влаги из воздуха гигроскопическим веществом;
 - б) неточный отсчет делений по шкале;
 - в) прикосновение цилиндра демпферной системы аналитических весов к оболочке;
 - г) неравноплечность коромысла весов;
 - д) испарение влаги при взвешивании влажных предметов;
 - е) невыверенные разновесы;
 - ж) несоответствие температуры взвешиваемого тела температуре воздуха весовой комнаты.
7. Указать способы уменьшения систематической погрешности:
 - а) при взвешивании;
 - б) при осаждении и промывании осадка;
 - в) при титровании с индикаторами.
8. Перечислить причины появления грубых погрешностей (промахов).
9. Дать классификацию погрешностей анализа по происхождению.
10. Перечислить инструментальные и методические погрешности.

11. Пояснить понятие «холостой» опыт. Как с его помощью можно оценить величину реактивной погрешности?
12. Дать определение абсолютной и относительной погрешностей.
13. Дать определение действительного значения определяемой величины.
14. Пояснить понятие правильности и воспроизводимости анализа.
15. Пояснить оценку правильности метода или методики анализа с помощью стандартного образца.
16. Описать приемы релятивизации и рандомизации погрешностей.
17. Привести примеры определения косвенных измерений.
18. Описать способ оценки достоверности единичных измерений.
19. Сформулировать правила вычисления суммарной предельной систематической погрешности результата анализа.
20. Пояснить определение количества значащих цифр в результатах вычислений при сложении и вычитании, умножении и делении, возведении в степень, при извлечении квадратного корня и логарифмировании.

2 Оценка случайных погрешностей с помощью методов математической статистики

2.1 Основные понятия классической статистики

Статистика — отрасль знаний, наука, в которой излагаются общие вопросы сбора, измерения и анализа массовых **статистических** (количественных или качественных) данных. Статистика дает математическое описание случайных процессов, например, влияния случайной погрешности на результаты химического анализа. Числовые значения, взятые для статистической обработки, называются **случайными величинами** и обозначаются x_1, x_2, \dots, x_n .

Классическая математическая статистика рассматривает идеализированный случай — бесконечно большое количество параллельных наблюдений и вводит понятия **генеральной совокупности** и **выборки** из нее. Всё абстрактно мыслимое количество случайных величин (от $-\infty$ до $+\infty$) носит название **генеральной совокупности**. Ограниченное множество случайных величин, эквивалентных друг другу со статистической точки зрения, например, результаты повторных анализов вещества, выполненных одним и тем же аналитическим методом, называется **выборкой** или **выборочной совокупностью**. Важнейшими характеристиками (параметрами) случайных величин является **среднее значение** и **стандартное отклонение**.

Различают характеристики генеральной совокупности, которые не зависят от числа наблюдений (μ, σ^2, σ), и выборочные характеристики случайных величин, которые зависят от числа наблюдений (\bar{x}, s^2, s).

Если между возможными значениями случайной величины x и плотностью вероятности их появления ($\varphi(x)$) можно установить определенное математическое соотношение, то оно называется **законом (или функцией) распределения случайной величины**. Закон распределения является важной характеристикой случайной величины.

Одной из основных задач химика-аналитика при оценке случайных погрешностей химического анализа является нахождение функции (закона) распределения, описывающей экспериментальные данные химического анализа. Случайная величина x считается заданной, если известна ее функция распределения $\varphi(x)$. Эта функция может быть представлена графически, в виде алгебраической зависимости или таблицы. Многочисленными исследованиями показано, что для большинства аналитических определений (при генеральной совокупности) результаты химического анализа подчиняются **закону нормального распределения (распределению Гаусса)**:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(x-\mu)^2 / 2\sigma^2} \quad (4)$$

где μ – **математическое ожидание** (среднее значение генеральной совокупности);

σ^2 – **дисперсия** (μ и σ^2 – параметры генеральной совокупности);

$\varphi(x)$ – плотность вероятности нормального распределения (величина, характеризующая вероятность появления отдельного измерения x_i).

Графическое изображение нормального распределения случайной величины x показано на рисунке 2.

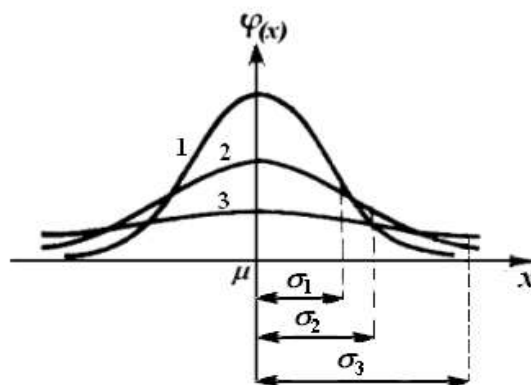


Рисунок 2 - Вид кривых плотности вероятности нормального распределения при разных параметрах σ

Дисперсия σ^2 (генеральная дисперсия) характеризует рассеяние случайных величин x_i относительно μ .

Положительное значение корня квадратного из дисперсии называют **генеральным стандартным отклонением** σ и также используют для характеристики рассеяния случайной величины x относительно μ в генеральной совокупности. Из кривой нормального распределения параметр σ определяется как половина расстояния между точками перегиба на кривой, а величина μ есть центр рассеяния случайных величин.

Закон нормального распределения для обработки результатов химического анализа применяют лишь в том случае, когда имеется большое количество данных ($n > 20-30$), при этом можно считать, что выборка достаточно полно представляет собой генеральную совокупность и вместо генеральных параметров μ и σ пользуются выборочными параметрами – средним значением выборки (\bar{x}) и выборочным стандартным отклонением (s).

Среднее значение выборки находят как среднеарифметическое из результатов отдельных измерений, то есть сумму случайных величин в выборке делят на число случайных величин:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum x_i}{n} . \quad (5)$$

Выборочное стандартное отклонение s большой выборки ($n > 20$) находят как корень квадратный из суммы квадратов отклонений между случайными величинами и средним значением выборки, разделенной на число случайных величин выборки (число параллельных результатов):

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}} . \quad (6)$$

Выборочные параметры \bar{x} и s (параметры выборочной совокупности) тем ближе к генеральным параметрам (μ и σ), чем больше выборка.

Если взять ряд выборок из одной генеральной совокупности и для каждой выборки определить среднее арифметическое значение \bar{x}_i , то получим ряд случайных величин $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n$, значения которых группируются вокруг своего математического ожидания μ , которое совпадает с математическим ожиданием единичных результатов анализа. Однако, рассеяние средних меньше, чем рассеяние единичных результатов. Дисперсия среднего арифметического ($V_{\bar{x}}$) в n раз меньше дисперсии единичного результата (V), а стандартное отклонение среднего ($s_{\bar{x}}$) в \sqrt{n} раз меньше стандартного отклонения единичного результата:

$$V_{\bar{x}} = \frac{V}{n}; \quad s_{\bar{x}} = \frac{s}{\sqrt{n}}.$$

Среднее значение \bar{x} , стандартное отклонение для единичного результата s и стандартное отклонение для среднего $s_{\bar{x}}$ имеют ту же размерность, что и измеряемая величина.

Данные химического анализа, как правило, подчиняются закону нормального распределения. Однако, если возникает сомнение в правомерности применения закона нормального распределения, то необходимо провести проверку распределения результатов анализа путем построения и анализа **гистограммы** (ступенчатой диаграммы) [3] (рисунки 3, 4).

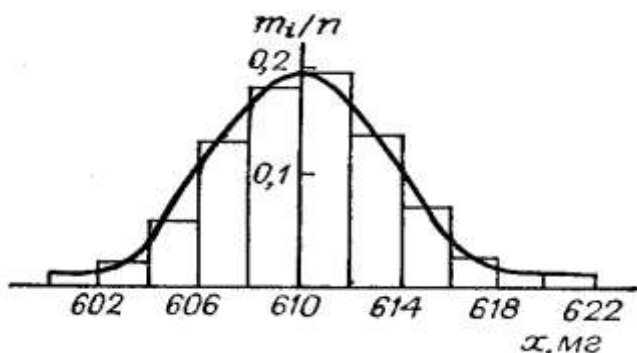


Рисунок 3 – Гистограмма результатов химического анализа

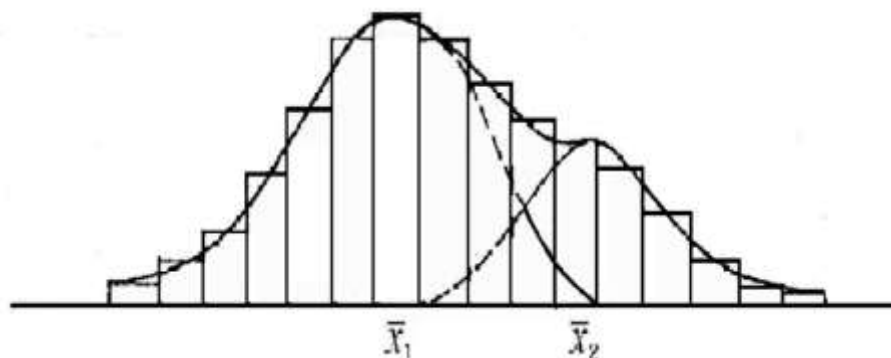


Рисунок 4 - Гистограмма результатов химического анализа при наличии систематической погрешности в части результатов

Если случайные величины x_i (результаты анализа) распределены нормально, то и их случайные погрешности ($\Delta x_i = x_i - \mu$) также распределены нормально.

При обработке данных химического анализа используют закон нормального распределения случайных погрешностей в нормированном виде, то есть случайную погрешность $(x_i - \mu)$ выражают в единицах стандартного отклонения:

$$U = \frac{x_i - \mu}{\sigma} = \frac{\Delta x}{\sigma}$$

и полученная случайная величина U также распределена нормально.

Аналогично нормируют и случайную погрешность среднего арифметического $\Delta \bar{x}$:

$$U_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\Delta \bar{x}}{\sigma} \cdot \sqrt{n}.$$

При многократном повторении измерений оценить случайную погрешность можно только с некоторой вероятностью, то есть, пользуясь статистическими методами, можно лишь указать вероятность появления случайной погрешности не больше некоторой величины. Поэтому оценку случайной погрешности проводят с помощью двух взаимосвязанных статистических параметров (критериев):

– **ширины доверительного интервала**, в пределах которого могут находиться результаты отдельных анализов;

– *доверительной вероятности* или *надежности* (P) того, что результаты отдельных анализов попадают в этот интервал.

Поскольку различают случайную погрешность единичного результата (Δx) и случайную погрешность среднего ($\bar{\Delta x}$), то доверительным интервалом и доверительной вероятностью характеризуют как результаты отдельных анализов, так и среднее значение. При статистической обработке данных многократного анализа для нахождения доверительной вероятности при заданном доверительном интервале (или наоборот) пользуются табулированными значениями интеграла $\Theta(U)$ (приложение А, таблица А.1),

$$\Theta(U) = \frac{1}{\sqrt{2\pi}} \int_0^U e^{-U^2/2} \cdot dU \quad . \quad (7)$$

который называют *нормированной функцией Лапласа (или Гаусса-Лапласа)*. Как видно из этой формулы, интеграл (а, следовательно, и доверительная вероятность), рассчитывается только для положительных значений нормированного доверительного интервала U , поскольку *нормированное нормальное распределение*, как и обычное распределение (кривая Гаусса), симметрично.

Для нахождения доверительной вероятности того, что случайная величина попадает в заданный интервал U , табличные значения вероятности надо увеличивать вдвое. Поскольку

$$U = \frac{x - \mu}{\sigma},$$

то, если случайная погрешность $|x - \mu|$ при многократном анализе не превышает $\pm 1\sigma$, $\pm 2\sigma$ и $\pm 3\sigma$ (что отвечает значению $U = \pm 1; \pm 2; \pm 3$), доверительная вероятность будет равна соответственно **0,6826** ($0,3413 \cdot 2$); **0,9544** ($0,4772 \cdot 2$) и **0,9973** ($0,49865 \cdot 2$). Эти три значения полезно запомнить.

Свойства кривой нормального распределения позволяют сделать выводы о вероятной величине случайной погрешности измерения при известном стандартном отклонении метода измерения.

Если величина σ известна (обычно стандартное отклонение s очень большой выборки принимают за величину, приближенную к σ), то можно утверждать, что 68,3 % всех единичных измерений имеют случайную погрешность не больше $\pm 1\sigma$; 95,4 % всех измерений имеют случайную погрешность не больше $\pm 2\sigma$; 99,7 % всех измерений имеют случайную погрешность не больше $\pm 3\sigma$.

Стандартное отклонение метода измерения является важным параметром для оценки величины случайных погрешностей параллельных измерений.

Оценками математического ожидания μ обычно служит *среднее арифметическое \bar{x} , а иногда среднее геометрическое x_g , медиана (срединное значение) \tilde{x} , или мода.*

Чтобы найти медиану, результаты измерений располагают в ряд (ранжированный ряд) в порядке возрастания их значений: $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$. Если n – нечетное число, то медиана \tilde{x} равна среднему члену ряда. Если n – четное число, то значение \tilde{x} равно полусумме двух средних членов ряда. Мода – наиболее часто встречающееся значение случайной величины при большой серии измерений.

Медиана в отличие от среднего арифметического нечувствительна к резко выделяющимся измерениям, поэтому при низком качестве измерений и малом числе n она является лучшей оценкой математического ожидания.

2.2 Применение функции нормированного распределения Гаусса-Лапласа для обработки результатов химического анализа

Большинство задач химической технологии в математической формулировке представляется системами уравнений, определяющих взаимосвязь многих факторов, от которых зависит течение процесса. Уравнения разделяются на линейные, алгебраические и трансцендентные.

В тех случаях, когда значение выборочного стандартного отклонения близко к генеральному, статистическая оценка результатов анализа не представляет трудностей.

Пример 2.1 - Среднее арифметическое значение из ряда определений массовой доли (%) диоксида кремния в образце силиката составляет 58,4; стандартное отклонение равно 0,6 %. Найти доверительную вероятность (P) того, что результат единичного (отдельного) анализа не выйдет за пределы $57,0 \% \leq x_i \leq 59,8 \%$.

Решение

1. Каковы допустимые интервалы отклонений от среднего значения ?

$$|x_1 - \bar{x}| = |\Delta x_1| = 58,4 - 57,0 = 1,4 \% ;$$

$$|x_1 - \bar{x}| = |\Delta x_1| = 59,8 - 58,4 = 1,4 \% .$$

Допустимые интервалы колебаний в обе стороны от среднего значения равны. Следовательно, равны и значения нормированного доверительного интервала U для единичного измерения.

2. Каково значение U ?

$$|U| = \frac{|\Delta x|}{\sigma} \approx \frac{|\Delta x|}{s} = \frac{1,4}{0,6} = 2,33 .$$

3. Каково значение функции Лапласа ?

Из таблицы приложения 1 для $U = 2,33$ функция Лапласа $\Theta = 0,4893$.

4. Какова доверительная вероятность того, что единичный результат анализа не выйдет за пределы от 57,0 до 59,8 ?

$$P = 0,4893 \cdot 2 = 0,9786 \approx 0,979.$$

Ответ: Доверительная вероятность (P) приблизительно равна 0,979 для $57,0 \% \leq w(\%)(\text{SiO}_2) \leq 59,8 \%$.

Пример 2.2 - Среднее арифметическое значение массовой доли (%) фосфора в чугуне по данным 8 параллельных анализов составляет 0,395. Генеральное стандартное отклонение равно 0,018 %. Какова доверительная вероятность того, что единичный результат и среднее арифметическое значение принадлежат интервалу от 0,370 до 0,420 % ?

Решение

1. Каковы допустимые интервалы отклонений от среднего значения?

$$|x_1 - \bar{x}| = |\Delta x_1| = 0,395 - 0,370 = 0,025;$$

$$|x_2 - \bar{x}| = |\Delta x_2| = 0,420 - 0,395 = 0,025 .$$

Допустимые интервалы отклонений в обе стороны от среднего значения равны. Следовательно, будут равны и значения нормированного доверительного интервала U .

2. Каково значение U для единичного результата анализа ?

$$|U| = \frac{|\Delta x|}{\sigma} = \frac{0,025}{0,018} = 1,39 .$$

3. Каково значение функции Лапласа для нормированного доверительного интервала единичного результата анализа?

Из таблицы приложения 1 для $U = 1,39$ функция Лапласа $\Theta = 0,4192$.

4. Какова доверительная вероятность того, что единичный результат анализа находится в интервале 0,370 ÷ 0,420?

$$P = 0,4192 \cdot 2 = 0,8384 \approx 0,838 .$$

5. Каково значение стандартного отклонения среднего значения?

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0,018}{\sqrt{8}} = 0,00636 .$$

6. Каково значение нормированного доверительного интервала для среднего значения ?

$$U_{\bar{x}} = \frac{|\Delta \bar{x}|}{\sigma_{\bar{x}}} = \frac{0,025}{0,00636} = 3,93 .$$

7. Каково значение функции Лапласа для нормированного доверительного интервала среднего ?

Из таблицы А.1 приложения А для $U = 3,93$ функция Лапласа $\Theta = 0,499950$.

8. Какова доверительная вероятность того, что среднее значение не выйдет за пределы от 0,370 до 0,420 (или какова вероятность того, что случайная погрешность среднего будет не больше 0,025 %) ?

$$P = 0,499950 \cdot 2 = 0,9999 .$$

Для среднего арифметического значения доверительная вероятность ($P = 0,9999$) значительно больше, чем для единичного результата анализа ($P = 0,838$).

Ответ: $P = 0,838$ для $0,370 \leq w_i(\%)(P) \leq 0,420$;

$P = 0,9999$ для $0,370 \leq \bar{w}(\%)(P) \leq 0,420$.

2.3 Статистика малых выборок. Применение распределения Стьюдента для обработки результатов химического анализа

В химическом анализе количество вещества в пробе устанавливают, проводя небольшое число параллельных определений ($2 \leq n < 20$). В этих случаях для оценки воспроизводимости и случайной погрешности пользуются методами математической статистики, разработанной для небольшого числа значений.

Результаты анализа при этом должны заведомо иметь нормальное распределение или близкое к нормальному.

В основе микростатистических оценок нормального распределения случайных величин лежит *распределение Стьюдента*³⁾, связывающее между собой три основные характеристики выборочной совокупности: ширину доверительного интервала, соответствующую ему доверительную вероятность и объем выборки n (или число степеней свободы f).

Результаты параллельных анализов рассматривают как случайную выборку из генеральной совокупности и вычисляют среднее значение и дисперсию малопредставительной выборки.

Среднее значение выборки – это сумма случайных величин в выборке, деленная на их число

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum x_i}{n} . \quad (8)$$

Дисперсия выборки (s^2) характеризует рассеяние результатов относительно среднего. При $n < 20$ дисперсию выборки находят по формуле

$$V = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} , \quad (9)$$

где $n - 1 = f$ – число степеней свободы.

Число степеней свободы $f = n - 1$, так как нахождение среднего исключает одно значение из числа независимых переменных n . Это значит, что рассеяние значений можно определить только при условии $n > 1$, при $n = 1$ понятие рассеяния

³⁾ Стьюдент – псевдоним английского математика и химика Госсета.

теряет смысл. Знаменатель уравнения для вычисления дисперсии *малопредставительной выборки* равен $f = n - 1$.

Для характеристики рассеяния результатов в выборочной совокупности используют также *стандартное отклонение* s и *относительное стандартное отклонение* s_r .

Стандартное отклонение (s) при $n < 20$ находят как квадратный корень из суммы квадратов отклонений между значениями случайных величин и средним значением выборки, разделенной на число степеней свободы случайных величин выборки

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} . \quad (10)$$

Данное выражение для стандартного отклонения неудобно для быстрого вычисления. Соотношения

$$s = \sqrt{\frac{\left(\sum_{i=1}^n x_i^2\right) - n(\bar{x})^2}{n - 1}} \quad \text{или} \quad s = \sqrt{\frac{\sum_{i=1}^n x_i^2 - \left[\left(\sum_{i=1}^n x_i\right)^2 / n\right]}{n - 1}}$$

абсолютно эквивалентны приведенному выше (10) и позволяют быстрее проводить расчет.

Относительное стандартное отклонение (s_r) может быть выражено в долях единицы или в процентах:

$$s_r = \frac{s}{\bar{x}} \quad \text{или} \quad s_r(\%) = \frac{s}{\bar{x}} \cdot 100 .$$

Относительное стандартное отклонение, выраженное в процентах, называется коэффициентом вариации или ковариацией.

Дисперсия, стандартное отклонение и относительное стандартное отклонение характеризуют воспроизводимость применяемого метода анализа.

Чем больше данных использовано при их вычислении, тем точнее полученные значения.

Часто для характеристики воспроизводимости метода используют величину стандартного отклонения s , а не дисперсию, поскольку стандартное отклонение выражается в тех же единицах, что и измеряемая величина.

Дисперсию среднего значения для малопредставительной выборки находят по формуле

$$V_{\bar{x}} = \frac{V}{n} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}, \quad (11)$$

а **стандартное отклонение для среднего значения** по формуле

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}}. \quad (12)$$

Распределение Стьюдента – это распределение нормированной случайной величины t . Если в статистике больших выборок рассматривают нормированную случайную величину $U_{\bar{x}}$ как меру отклонения среднего результата анализа от математического ожидания μ (в генеральном стандартном отклонении среднего - $\sigma_{\bar{x}}$), то $t_{\bar{x}}$ – аналогичная мера в стандартном отклонении для оценки среднего результата ($s_{\bar{x}}$) в случайной выборке.

$$U_{\bar{x}} = \frac{|\bar{x} - \mu|}{\sigma_{\bar{x}}} = \frac{\Delta \bar{x} \cdot \sqrt{n}}{\sigma}; \quad t_{\bar{x}} = \frac{|\bar{x} - \mu|}{s_{\bar{x}}} = \frac{\Delta \bar{x} \cdot \sqrt{n}}{s}. \quad (13)$$

Сумма квадратов случайных величин $U_{x_i} (\chi^2 = \sum_{i=1}^n U_{x_i}^2)$ имеет свой закон распределения, который называется χ^2 -распределением (хи-квадрат) с числом степеней свободы $f = n - 1$. Можно сравнить величины U и t , нормированные соответственно в единицах генерального и выборочного стандартного отклонения для *единичного* результата:

$$U_{x_i} = \frac{|x_i - \mu|}{\sigma} = \frac{\Delta x}{\sigma}, \quad t_{x_i} = \frac{|x_i - \mu|}{s} = \frac{\Delta x}{s},$$

здесь $t_x = t_{x_i}$ и обозначается сокращенно t .

Величины U и t не равны между собой, поскольку $\sigma \neq s$, но $t \rightarrow U$ так же, как и $s \rightarrow \sigma$ при $n \rightarrow \infty$.

В отличие от U коэффициент Стьюдента t зависит не только от заданной доверительной вероятности, но и от числа степеней свободы, которое используется для вычисления стандартного отклонения.

Аналитическое выражение распределения Стьюдента $\varphi(t)$ (плотности вероятности t -распределения) громоздко, поэтому его представляют в табличном (приложение Б, таблица Б.1) или в графическом виде. Значение t (коэффициента Стьюдента) зависит как от заданной доверительной вероятности P , так и от числа параллельных определений n . В химическом анализе, как правило, задают доверительную вероятность 0,95.

Вид кривых плотности вероятности $\varphi(t)$ для трех значений степеней свободы (f) приведен на рисунке 5.

Для $f \rightarrow \infty$ кривая $\varphi(t)$ совпадает с кривой нормированного стандартного распределения $\varphi(U)$. Для малых выборок кривая $\varphi(t)$ идет более полого, медленнее приближаясь к оси абсцисс при больших значениях $|t|$.

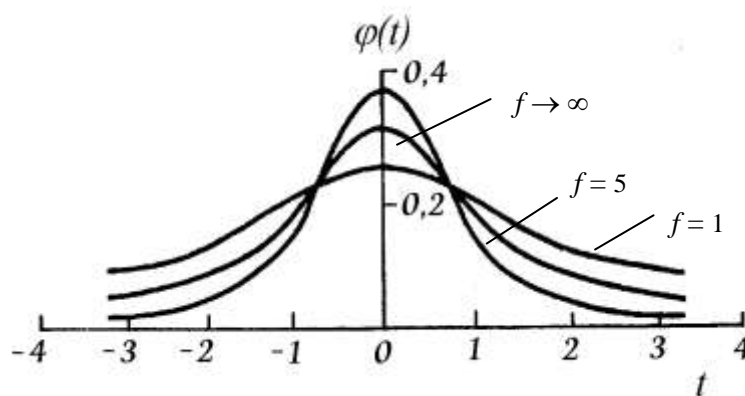


Рисунок 5 - Зависимость плотности вероятности $\varphi(t)$ от ширины доверительного интервала t в распределении Стьюдента при различном числе степеней свободы f

Коэффициент эксцесса (коэффициент островершинности) — мера остроты пика распределения случайной величины. Коэффициент эксцесса при строго нормальном распределении равен нулю. Если его значение больше нуля, то вершина распределения более заостренная по сравнению с нормальным распределением (такое распределение получают, например, если результаты измерений отбирают преднамеренно), а если меньше нуля — вершина распределения более пологая.

При одинаковой ширине доверительного интервала доверительная вероятность, оцененная по Стьюденту, всегда меньше доверительной вероятности, оцененной по нормированному распределению Гаусса-Лапласа. При одинаковой доверительной вероятности ширина доверительного интервала по Стьюденту всегда больше чем по распределению Гаусса-Лапласа. Другими словами, оценка по Стьюденту путем расширения доверительного интервала учитывает неполноту статистической выборки.

При обработке результатов химического анализа рассчитывают среднее значение \bar{x} и воспроизводимость метода анализа дисперсией V , стандартным отклонением s или относительным стандартным отклонением s_r . Кроме того, результаты анализа характеризуют шириной доверительного интервала для среднего значения $(\pm \Delta \bar{x})$, ограничивающего область, в которой при заданной

доверительной вероятности (и при отсутствии систематических погрешностей) лежит действительное значение. Этот интервал можно рассчитать, пользуясь формулой:

$$t = \frac{\Delta \bar{x} \cdot \sqrt{n}}{s}, \quad \text{откуда} \quad \Delta \bar{x} = \pm \frac{t \cdot s}{\sqrt{n}}.$$

Достоверность единичного результата оценивают по доверительному интервалу для единичного результата по формуле:

$$\Delta x = \pm t \cdot s.$$

Пример 2.3- При определении оптической плотности растворов комплекса катионов никеля с диметилглиоксимом при постоянной концентрации Ni^{2+} -ионов были получены следующие значения оптической плотности: 0,292; 0,294; 0,290; 0,290; 0,295. Найти границы доверительного интервала для среднего значения оптической плотности раствора при доверительной вероятности $P = 0,95$ и относительную погрешность определения среднего (%).

Решение

1. Обозначим значение оптической плотности за x и найдем среднее значение:

$$\bar{x} = \frac{0,292 + 0,294 + 0,290 + 0,290 + 0,295}{5} = 0,292.$$

2. Какова величина стандартного отклонения для единичного измерения ?

$$s = \sqrt{\frac{(0,292 - 0,292)^2 + (0,294 - 0,292)^2 + 2 \cdot (0,290 - 0,292)^2 + (0,295 - 0,292)^2}{4}} = 0,0023.$$

3. Какова ширина доверительного интервала для среднего значения (случайная погрешность среднего) ?

$$\pm \Delta \bar{x} = \frac{t \cdot s}{\sqrt{n}} = \frac{2,78 \cdot 0,0023}{\sqrt{5}} = 0,00286 \approx 0,003.$$

4. Какова относительная погрешность определения среднего (%) ?

$$\pm \frac{\Delta \bar{x}}{\bar{x}} \cdot 100 = \pm \frac{0,00286}{0,292} \cdot 100 = \pm 0,979 \approx \pm 0,98.$$

5. Каковы интервальные значения измеряемой величины?

$$\bar{x} + \Delta \bar{x} = 0,292 + 0,003 = 0,295.$$

$$\bar{x} - \Delta \bar{x} = 0,292 - 0,003 = 0,289.$$

Ответ: Истинное значение оптической плотности с вероятностью 0,95 будет находиться в интервале от 0,289 до 0,295; относительная ошибка определения среднего составляет $\pm 0,98 \%$.

В рассмотренном примере статистически обрабатываются не результаты анализа, а измеренные значения показаний прибора.

Результаты вычислений не могут быть выражены точнее, чем те, что использованы для статистической обработки. Расчет \bar{x} , s , $\Delta \bar{x}$ проводят всегда с сохранением одной лишней цифры, в сравнении с числом цифр, использованных для обработки данных. В конечном результате эту лишнюю цифру округляют. Результат измерения и его погрешность должны быть выражены числами с одинаковым количеством знаков после запятой.

Экспериментальные данные и результаты статистической обработки рекомендуется представлять в виде таблицы (таблица 3):

Таблица 3 - Экспериментальные данные и статистическая обработка фотометрического определения оптической плотности раствора Ni^{2+} -ионов с диметилглиоксимом

x_i	n	\bar{x}	s	$\frac{t \cdot s}{\sqrt{n}}$	$\bar{x} \pm \frac{t \cdot s}{\sqrt{n}}$
0,292	5	0,292	0,0023	0,003	0,289 ÷ 0,295
0,294					
0,290					
0,290					
0,295					

Пример 2.4 - При определении массовой доли (%) никеля в стали среднее из восьми значений равно 1,76. Выборочное стандартное отклонение равно 0,08 %. Найти коэффициент вариации и границы доверительного интервала для среднего и для единичного результата анализа при $P = 0,95$.

Решение

1. Каково значение коэффициента вариации ?

$$s_r(\%) = \frac{s}{\bar{x}} \cdot 100 = \frac{0,08}{1,76} \cdot 100 = 4,54 \text{ .}$$

2. Каково значение коэффициента Стьюдента t при $P = 0,95$ и $n = 8$?

Из таблицы Б.1 приложения Б следует, что $t_{P,f} = t_{0,95;7} = 2,37$.

3. Какова ширина доверительного интервала для среднего значения ?

$$\pm \Delta \bar{x} = \frac{t \cdot s}{\sqrt{n}} = \frac{2,37 \cdot 0,08}{\sqrt{8}} = 0,067 \approx 0,07(\%).$$

4. Каковы граничные значения среднего ?

$$\bar{x} + \Delta \bar{x} = 1,76 + 0,07 = 1,83 \text{ (\%).}$$

$$\bar{x} - \Delta \bar{x} = 1,76 - 0,07 = 1,69 \text{ (\%).}$$

5. Какова ширина доверительного интервала для единичного результата ?

$$\pm \Delta x_i = t \cdot s = 2,37 \cdot 0,08 = 0,19 \text{ .}$$

6. Каковы граничные значения для единичного результата ?

$$\bar{x} + \Delta x_i = 1,76 + 0,19 = 1,95 \text{ (\%).}$$

$$\bar{x} - \Delta x_i = 1,76 - 0,19 = 1,57 \text{ (\%).}$$

Ответ: $1,69 \leq \bar{x} \leq 1,83$; $1,57 \leq x_i \leq 1,95$.

Расчет среднего значения \bar{x} , дисперсии V , стандартного отклонения для среднего результата $s_{\bar{x}}$, доверительного интервала для среднего $|\Delta \bar{x}|$ и интервальных границ среднего $\bar{x} \mp \Delta \bar{x}$ можно выполнить с помощью компьютера.

2.4 Проверка значений измерений на наличие грубых погрешностей (промахов)

Методы математической статистики позволяют провести проверку результатов на наличие грубых погрешностей (промахов) по разным критериям. Самым простейшим из них является **критерий $3s$** . Грубой погрешностью считается результат, отклонение которого от среднего ($|x_i - \bar{x}|$) выходит за пределы величины $3s$. В расчет среднего значения и стандартного отклонения s включается подозреваемый на промах результат x_i , который, на наш взгляд, недопустимо мал или велик. Если неравенство $|x_i - \bar{x}| > 3s$ будет иметь место для нескольких значений, то отбрасывают сначала лишь то значение, для которого отклонение единичного результата от среднего $|x_i - \bar{x}|$ оказалось наибольшим. Из оставшихся значений x_i вычисляют новые значения среднего \bar{x} , стандартного отклонения s и повторяют проверку результатов на грубую погрешность.

Проверку результатов на грубую погрешность можно провести без вычисления стандартного отклонения, используя Q -критерий. Используя результаты анализов, $Q_{расч.}$ вычисляют по формуле:

$$Q_{расч.} = \frac{x_i - x_c}{R}, \quad (14)$$

где x_i – результат, подозреваемый на промах;

x_c – результат, соседний по значению;

$R = x_{max} - x_{min}$ – размах варьирования или диапазон выборки (разность между предельными значениями определяемой величины).

Величину $Q_{расч.}$ сопоставляют с табличным значением ($Q_{табл.}$) для данного n и доверительной вероятности P (таблица Г.1, приложение Г). Наличие грубой погрешности доказано, если $Q_{расч.} > Q_{табл.}$. Если сомнение вызывает не одно, а

несколько значений, то порядок их проверки на грубую погрешность такой же, как для критерия $3s$.

В результате выбраковки подозреваемых на промах результатов может быть исключено не более одной трети всех значений. Если исключению подлежит большее число значений, все полученные результаты анализа считают неудовлетворительными и анализы повторяют.

2.4.1 Проверка наличия систематической погрешности по результатам анализа стандартного образца

По результатам анализа стандартных образцов, в которых содержание определяемого элемента считается действительным, можно судить о наличии или отсутствии систематической погрешности в определении содержания этого элемента. Если действительное значение содержания определяемого элемента находится в пределах доверительного интервала среднего значения, то считают, что отличие среднего результата от действительного значения обусловлено случайными погрешностями, а систематической погрешности нет.

Если расхождение между действительным (μ) и средним значением больше или равно ширине доверительного интервала ($|\bar{x} - \mu| \geq |\Delta\bar{x}|$), то говорят о наличии систематической погрешности. В этом случае необходимо выяснить причину появления систематической погрешности и устранить ее.

Выражение $|\bar{x} - \mu| \geq |\Delta\bar{x}|$ можно представить в виде:

$$|\bar{x} - \mu| \geq \left| \frac{t \cdot s}{\sqrt{n}} \right|.$$

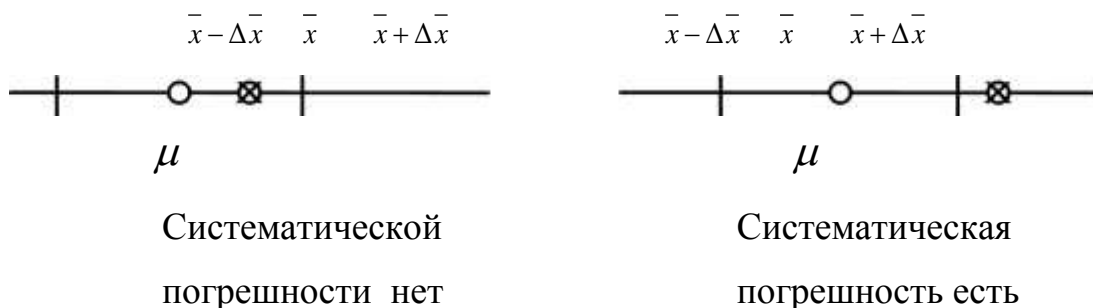
Если существует надежное значение стандартного отклонения s , то можно его принять за генеральное стандартное отклонение σ и в последнем выражении

заменить s на σ , а коэффициент Стьюдента t на U (доверительный интервал, выраженный в единицах генерального стандартного отклонения):

$$|\bar{x} - \mu| \geq \left| \frac{U \cdot \sigma}{\sqrt{n}} \right|.$$

Величину U находят из таблицы приложения А. Например, если $P = 0,95$, то функция Лапласа $\Theta = 0,475$, а $U \approx 1,96$.

Ниже схематично показано отсутствие или наличие систематической погрешности в результатах анализа стандартного образца.



Пример 2.5 - При определении содержания (мкг/дм^3) ионов меди в стандартном образце были получены следующие данные : 5,1; 5,5; 5,4; 5,8; 5,2; 7,1. Действительное значение содержания ионов меди равно $5,3 \text{ мкг/дм}^3$. Провести статистическую обработку полученных данных.

Решение

1. Проверяем наличие грубых погрешностей по Q -критерию. Предполагаем, что значение 7,1 является грубой погрешностью. Рассчитываем значение Q -критерия по результатам анализов ($Q_{расч}$):

$$Q_{расч.} = \frac{x_n - x_c}{x_{\max} - x_{\min}}; \quad Q_{расч} = \frac{7,1 - 5,8}{7,1 - 5,1} = 0,65.$$

2. Находим значение Q -критерия ($Q_{табл.}$) из таблицы приложения Г: при $n = 6$ и $P = 0,95$. $Q_{табл.} = 0,56$.

3. Сравниваем: $0,65 > 0,56$, то есть $Q_{расч} > Q_{табл.}$, поэтому значение 7,1 считаем недостоверным и исключаем из статистически обрабатываемых значений.

4. Вычисляем среднее значение, стандартное отклонение ⁴⁾, доверительный интервал и интервальные значения для среднего.

5. Результаты статистической обработки сводим в таблицу 4.

Таблица 4 - Результаты статистической обработки определения содержания ионов меди в образце

x_i	n	\bar{x}	s	$\frac{t \cdot s}{\sqrt{n}}$	$\bar{x} \pm \frac{t \cdot s}{\sqrt{n}}$
5,1	5	5,4	0,27	0,3	5,1÷5,7
5,5					
5,4					
5,8					
5,2					

Для доверительной вероятности ($P = 0,95$) интервальные значения для среднего результата (от 5,1 до 5,7) в 95 % случаев среднее значение результата не выйдет за эти пределы.

6. Использованный метод определения ионов меди не содержит систематической погрешности, так как действительное содержание ($\mu = 5,3 \text{ мкг/дм}^3$) ионов меди в анализируемом образце не выходит за пределы интервальных значений для среднего результата.

7. Находим относительную погрешность среднего результата с доверительной вероятностью 0,95:

$$\pm \frac{\Delta \bar{x} \cdot 100}{\mu} = \pm \frac{0,3 \cdot 100}{5,3} = \pm 5,3\%.$$

Ответ: $\bar{x} = 5,4 \text{ мкг/дм}^3$; $\bar{x} \pm \Delta \bar{x} = 5,1 \div 5,7$; систематическая погрешность отсутствует.

⁴⁾ Целесообразно стандартное отклонение s определять предварительно. В этом случае можно получить величину s , близкую к теоретической.

Если для анализируемого образца, действительное содержание неизвестно, то относительную погрешность среднего результата находят по формуле:

$$\pm \frac{\Delta \bar{x} \cdot 100}{\bar{x}} .$$

Можно сформулировать **основные этапы обработки результатов химического анализа:**

- 1) проверка результатов на наличие грубых погрешностей по Q -критерию;
- 2) нахождение среднего значения и (\bar{x}) из достоверных данных параллельных определений;
- 3) нахождение дисперсии (V) результатов анализов;
- 4) нахождение стандартного отклонения для единичного результата (s);
- 5) нахождение (при необходимости) относительного стандартного отклонения (s_r или $s_r(\%)$);
- 6) нахождение табличного значения коэффициента Стьюдента ($t_{P,f}$) при заданной доверительной вероятности (P) и известном числе степеней свободы (f) (приложение Б, таблица Б.1);
- 7) нахождение ширины доверительного интервала для среднего значения ($\pm \Delta \bar{x}$) (абсолютной погрешности определения);
- 8) нахождение интервальных границ среднего результата ($\bar{x} \pm \Delta \bar{x}$);
- 9) нахождение относительной погрешности среднего ($\pm \frac{\Delta \bar{x}}{\bar{x}} \cdot 100$);
- 10) проверка наличия систематической погрешности в определении.

2.5 Статистические методы как средства проверки моделей

Большинство исследований в химии проводятся с целью проверки известных и вновь разрабатываемых моделей. Если результаты экспериментов не

подтверждают модель, ее отвергают и предлагают новую. Если наблюдается соответствие между экспериментальными и ожидаемыми результатами, такая модель может быть основой для последующих экспериментов. Если в дальнейшем она подтверждается достаточным количеством экспериментальных данных, модель признается до тех пор, пока другие исследования не опровергнут ее.

Экспериментальные данные редко совпадают с модельными расчетами, так как результаты анализа и производные от них величины содержат неустранимую случайную погрешность. Исследователю часто приходится решать вопрос: является ли это расхождение признаком ошибочности расчетов или результатом случайных погрешностей измерений.

При заданной доверительной вероятности P и соответствующему ей риску статистические методы позволяют дать объективную интерпретацию результатов анализа, β - уровень значимости - показывает вероятность выхода случайной величины за границы доверительного интервала ($\beta = 1 - P$).

К таким статистическим проверкам относится нуль-гипотеза, основанная на сравнении численных значений двух разных величин, которые гипотетически считают равными. Используя распределение Фишера, Стьюдента, или другие распределения, по результатам эксперимента оценивают вероятность расхождения между двумя сравниваемыми величинами вследствие случайных погрешностей или рассчитывают критическое значение некоторой случайной величины ($t_{крит.}$; $F_{крит.}$). Если расхождение равно или больше расхождения $\beta = 0,05$, то оно считается значимым и нуль-гипотеза не принимается.

Чаще всего химики проводят такие проверки:

- сравнение дисперсий s_1^2 и s_2^2 двух серий параллельных измерений (анализов);
- сравнение средних результатов двух выборок \bar{x}_1 и \bar{x}_2 ;
- сравнение стандартного отклонения s малой выборки и стандартного отклонения σ большой выборки;

– сравнение среднего результата анализа \bar{x} и величины μ , принимаемой за истинное значение. При этом нуль-гипотеза основывается на предположении о незначимости различия между \bar{x} и μ .

2.5.1 Сравнение двух методов по воспроизводимости. Критерий Фишера

В практике химического анализа часто приходится сравнивать методы по воспроизводимости. Например, при анализе одной и та же пробы двумя разными методами получены две серии результатов (две выборочные совокупности), для которых рассчитаны значения выборочных дисперсий V_1 и V_2 . Очевидно, что выборочные дисперсии не совпадают. Необходимо выяснить, является ли различие дисперсий значимым (неслучайным), когда один из методов дает более воспроизводимые результаты, чем другой, или различие дисперсий имеет случайный характер. Эта задача решается с помощью F -критерия. Проверяют нуль-гипотезу о равенстве генеральных дисперсий σ_1^2 и σ_2^2 , если известны выборочные дисперсии s_1^2 и s_2^2 . Критические значения критерия Фишера ($F_{крит.}$) приведены в таблице Д.1, приложения Д для уровня значимости $\beta = 0,05$.

Случайные величины $F_{крит.}$ ($F_{крит.} = \frac{s_1^2}{s_2^2} = \frac{V_1}{V_2}$) (отношение выборочных дисперсий)

подчиняются закону распределения Фишера. Функция плотности вероятности $\varphi(F)$ имеет два параметра – f_1 и f_2 (числа степеней свободы дисперсий s_1^2 и s_2^2 соответственно).

F -критерий позволяет при заданном уровне определить, случайно ли различие двух дисперсий или значимо. Для этого рассчитывают коэффициент $F_{эсп.}$ как отношение V_1/V_2 (условно принято в числитель ставить большее число, чтобы величина $F_{эсп.}$ была больше единицы:

$$F_{\text{эсп.}} = \frac{s_1^2}{s_2^2} \quad (\text{при } s_1^2 > s_2^2).$$

Полученное значение $F_{\text{эсп.}}$ сравнивают с $F_{\text{табл.}}$ (с $F_{\text{крит.}}$) при числе степеней свободы $f_1 = n_1 - 1$ и $f_2 = n_2 - 1$. В таблице Д.1 приложения Д число степеней свободы большей дисперсии приводится в горизонтальной строке, а меньшей дисперсии – в вертикальном столбце. Если $F_{\text{эсп.}} \geq F_{\text{табл.}}$, то различие двух дисперсий значимо, то есть методы химического анализа различаются по воспроизводимости.

Пример 2.6 - При определении массовой доли (%) хрома в одной и той же пробе стали двумя разными методами химического анализа получены следующие результаты:

Номер метода химического анализа	Среднее значение \bar{x}	Число степеней свободы $f = n-1$	Дисперсия s^2
1	0,94	3	$4,2 \cdot 10^{-3}$
2	0,92	4	$7,7 \cdot 10^{-4}$

Проверить, значимо ли различаются дисперсии, если $\beta=0,05$.

Решение

1. Рассчитываем $F_{\text{эсп.}}$:

$$F_{\text{эсп.}} = \frac{s_1^2}{s_2^2} = \frac{4,2 \cdot 10^{-3}}{7,7 \cdot 10^{-4}} = 5,45.$$

2. Находим $F_{\text{табл.}}$ для $f_1=3$ и $f_2=4$:

$$F_{\text{табл.}} = 6,6.$$

3. Сравниваем $F_{\text{эсп.}}$ и $F_{\text{табл.}}$:

$$5,45 < 6,6; \quad F_{\text{эсп.}} < F_{\text{табл.}},$$

следовательно, расхождение между дисперсиями незначимо.

Ответ: Нельзя утверждать, что два сравниваемых метода имеют разную воспроизводимость.

Аналогично можно сравнивать две методики химического анализа по воспроизводимости, сравнивать дисперсии результатов анализа, полученных в разных лабораториях, на разных приборах, разными химиками-аналитиками, выполненных в разное время.

2.5.2. Сопоставление средних результатов двух химических анализов. Критерий Фишера

Часто возникает необходимость сравнивать средние значения двух выборочных совокупностей с числом параллельных анализов n_1 и n_2 . По результатам анализов найдены средние значения \bar{x}_1 и \bar{x}_2 , которые не совпадают. Обусловлено ли отличие в средних значениях случайными погрешностями или существованием систематических погрешностей? Для решения этой задачи нуль-гипотеза формулируется следующим образом: равна ли нулю разность $\mu_1 - \mu_2$, если $\bar{x}_1 \neq \bar{x}_2$?

Ситуации могут быть разными: одна и та же проба анализируется разными аналитиками, или в разных лабораториях, или разными методами. Может быть и такая ситуация: анализируются разные пробы, но предполагается, что они идентичные. Нуль-гипотезу можно проверить с помощью t -критерия Стьюдента, если есть основание полагать, что:

- в каждом случае результаты распределены нормально;
- дисперсии двух серий наблюдений s_1^2 и s_2^2 не отличаются значимо одна от другой.

Значимость отличия дисперсий проверяется с помощью F -критерия, как это было показано в разделе 2.5.1. Если отличие дисперсий значимо, то сравнивать средние значения нельзя. Если указанные выше два условия выполняются, то

рассчитывают средневзвешенную дисперсию (среднее взвешенное двух дисперсий) по формуле:

$$(\bar{s})^2 = \frac{f_1 s_1^2 + f_2 s_2^2}{f_1 + f_2} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}. \quad (15)$$

Затем рассчитывают коэффициент Стьюдента по формуле

$$t_{\text{экс.}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{(\bar{s})^2}} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}, \quad (16)$$

сравнивают $t_{\text{экс.}}$ с $t_{\text{табл.}}$ для выбранной доверительной вероятности (как правило, для $P = 0,99$) при $f = f_1 + f_2 = n_1 + n_2 - 2$. Если $t_{\text{экс.}} \geq t_{\text{табл.}}$, то расхождение между двумя средними значимо (неслучайно), выборки не принадлежат одной генеральной совокупности и $\mu_1 \neq \mu_2$. Если речь идет об идентичности проб, то можно сказать, что пробы не идентичны. Если сравнивать два метода или работу двух аналитиков, то можно сказать, что по крайней мере в одном из двух средних значений, а, возможно, и в обоих, есть систематическая погрешность. Если $t_{\text{экс.}} < t_{\text{табл.}}$, то нуль-гипотеза $\mu_1 - \mu_2 = 0$ принимается и можно все данные рассматривать как единую выборочную совокупность в $(n_1 + n_2)$ результатов.

Пример 2.7 - Одна и та же проба железорудного концентрата проанализирована в двух лабораториях одним и тем же методом на содержание железа. Получены следующие значения массовой доли железа (%):

Номер лаборатории	Число параллельных определений	Среднее арифметическое	Стандартное отклонение
№	n	\bar{x}	s
1	6	21,3	0,40
2	5	20,8	0,28

Совместны ли результаты одной и той же пробы, полученные в двух лабораториях ($\beta=0,01$) ?

Решение

1. Проверяем, значимо ли различаются дисперсии?

$$F_{\text{экс.}} = \frac{s_1^2}{s_2^2} = \frac{(0,40)^2}{(0,28)^2} = 2,04.$$

$$F_{\text{табл.}} = 6,3 \text{ (для } f_1=5; f_2=4 \text{ и } \beta=0,05).$$

$$2,04 < 6,3; F_{\text{экс.}} < F_{\text{табл.}}$$

Следовательно, различие дисперсий незначимо и можно сравнивать средние значения.

2. Рассчитываем средневзвешенную дисперсию:

$$(\bar{s})^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} = \frac{5 \cdot (0,40)^2 + 4 \cdot (0,28)^2}{6 + 5 - 2} = 0,12.$$

3. Рассчитываем коэффициент Стьюдента:

$$t_{\text{экс.}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\bar{s}} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = \frac{(21,3 - 20,8)}{0,352} \cdot \sqrt{\frac{6 \cdot 5}{11}} = 2,35.$$

4. Находим из таблицы Б.1 приложения Б коэффициент Стьюдента:

$$t_{\text{табл.}} = 3,25 \text{ (при } P=0,99, f=9).$$

5. Сравниваем $t_{\text{экс.}}$ с $t_{\text{табл.}}$:

$$2,35 < 3,25; t_{\text{экс.}} < t_{\text{табл.}}$$

Следовательно, расхождение между средними значениями незначимо и обусловлено случайными погрешностями. Можно определять новое среднее (\bar{x}) единой выборочной совокупности и дисперсию (V):

$$\bar{x} = \frac{\sum x_i}{n_1 + n_2}; \quad s^2 = \frac{\sum (\bar{x} - x_i)^2}{n_1 + n_2 - 1}.$$

Ответ: Нуль-гипотеза принимается; результаты совместны.

2.6 Вычисление суммарной случайной погрешности результата химического анализа

При вычислениях, содержащих несколько последовательных арифметических действий случайную погрешность конечного результата химического анализа можно оценить через случайные погрешности отдельных этапов. Для этого обычно ограничиваются только расчетом абсолютного или относительного стандартного отклонения (s), поэтому оно должно быть известно для каждой метрологической операции, проведенной при анализе.

Задачу можно сформулировать таким образом. Пусть y – результат косвенного химико-аналитического определения; x_1, x_2, \dots, x_n – аргументы, значения которых измеряются в ходе химического анализа, а f – конкретная функциональная зависимость, которая их связывает:

$$y = f(x_1, x_2, \dots, x_n).$$

Чему равна случайная погрешность в определении величины y , если известны случайные погрешности, допускаемые при измерении отдельных значений x_i ?

Для оценки случайных погрешностей косвенных измерений используется математический аппарат дифференциального исчисления. С его помощью выведены правила для нахождения абсолютного и относительного стандартного отклонения для разных функций.

Правило 1. Для простейших функций двух аргументов

$$y = x_1 + x_2 \quad \text{или} \quad y = x_1 - x_2$$

абсолютное стандартное отклонение суммы (или разности) равно квадратному корню из суммы квадратов **абсолютных** стандартных отклонений:

$$s_y = \sqrt{s_{x_1}^2 + s_{x_2}^2}.$$

Правило 2. Для функции вида $y = x_1 \cdot x_2$ или $y = x_1/x_2$ имеем:

$$\left| \frac{s_y}{y} \right| = \sqrt{\left(\frac{s_{x_1}}{x_1} \right)^2 + \left(\frac{s_{x_2}}{x_2} \right)^2}.$$

Относительное стандартное отклонение произведения (или частного от деления) равно квадратному корню из суммы квадратов **относительных** стандартных отклонений сомножителей (или делимого и делителя)⁵⁾.

Для функции вида $y = (x_1 \cdot x_2 \cdot \dots \cdot x_k) / (x_l \cdot x_m \cdot x_n)$, часто встречающейся при расчетах в титриметрическом анализе, имеем

$$\left| \frac{s_y}{y} \right| = \sqrt{\sum_{i=1}^n \left(\frac{s_{x_i}}{x_i} \right)^2}.$$

Правило 3. Для степенной функции вида $y = x^n$, где n – число, которое не содержит погрешность, имеем:

$$\left| \frac{s_y}{y} \right| = \left| n \cdot \frac{s_x}{x} \right|.$$

Относительное стандартное отклонение степенной функции равно **относительному** стандартному отклонению величины, которая возводится в степень, умноженному на показатель степени.

Правило 4. Для функции вида $y = \lg x$, имеем:

$$s_y = 0,4343 \cdot \left| \frac{s_x}{x} \right|.$$

Абсолютное стандартное отклонение логарифма (по основанию 10) равно **относительному** стандартному отклонению логарифмируемой величины, умноженному на 0,4343.

При вычислении антилогарифма $x = \text{antlg } y$

⁵⁾ При вычислении относительных стандартных отклонений вместо неизвестных истинных значений y и x_i используют средние значения \bar{y} и \bar{x}_i .

$$\left| \frac{s_x}{x} \right| = \frac{s_y}{0,4343} \cdot$$

Формулы для нахождения суммарной случайной погрешности химического анализа представлены в таблице 5.

Случайная погрешность суммы или разности определяется абсолютными величинами случайных погрешностей. Погрешность произведения, частного от деления, возведения в степень – относительными величинами. Определив относительную погрешность конечного результата анализа, можно рассчитать его абсолютную погрешность и, наоборот, зная абсолютную погрешность, можно найти относительную погрешность.

Таблица 5 - Формулы для нахождения суммарной случайной погрешности косвенных измерений

Функция	Формула	Функция	Формула
$y = x_1 + x_2$	$s_y = \sqrt{s_{x_1}^2 + s_{x_2}^2}$	$y = x^n$	$\left \frac{s_y}{y} \right = \left n \cdot \frac{s_x}{x} \right $
$y = x_1 - x_2$	$s_y = \sqrt{s_{x_1}^2 + s_{x_2}^2}$	$y = \ln x$	$s_y = \left \frac{s_x}{x} \right $
$y = x_1 \cdot x_2$	$\left \frac{s_y}{y} \right = \sqrt{\left(\frac{s_{x_1}}{x_1} \right)^2 + \left(\frac{s_{x_2}}{x_2} \right)^2}$	$y = \lg x$	$s_y = 0,4343 \cdot \left \frac{s_x}{x} \right $
$y = x_1/x_2$	$\left \frac{s_y}{y} \right = \sqrt{\left(\frac{s_{x_1}}{x_1} \right)^2 + \left(\frac{s_{x_2}}{x_2} \right)^2}$	$x = \text{antlg } y$	$\left \frac{s_x}{x} \right = \frac{s_y}{0,4343}$

Пример 2.8 - При определении массовой доли (%) железа в руде на титрование пяти одинаковых по объему проб раствора, содержащего Fe^{2+} -ионы, израсходовано 6,60; 6,63; 6,61; 6,59; 6,64 см^3 раствора дихромата калия. Вычислите массовую долю (%) железа в руде и абсолютное стандартное отклонение результата анализа, если объем пипетки $9,95(\pm 1 \cdot 10^{-2}) \text{ см}^3$; молярная концентрация эквивалента дихромата калия $0,04959(\pm 9 \cdot 10^{-5}) \text{ моль/дм}^3$; объем колбы

200,0(±0,2)см³, масса навески 0,6601(±4·10⁻⁴) г (в скобках указаны абсолютные стандартные отклонения⁶⁾ разных метрологических операций).

Решение

1. Каково среднее значение объема раствора дихромата калия ?

$$\bar{V}(\text{K}_2\text{Cr}_2\text{O}_7) = \frac{6,60 + 6,63 + 6,61 + 6,59 + 6,69}{5} = 6,614 \text{ (см}^3\text{)}.$$

2. Каково стандартное отклонение измерения объема бюреткой?

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(6,60 - 6,614)^2 + (6,63 - 6,614)^2 + 0 + (6,59 - 6,614)^2 + (6,64 - 6,614)^2}{5-1}} = 2,12 \cdot 10^{-2} \text{ (см}^3\text{)}$$

3. Какова массовая доля железа в руде?

Массовую долю железа в руде рассчитываем по формуле:

$$w(\%)(\text{Fe}) = \frac{c(f_{\text{экв.}}(\text{K}_2\text{Cr}_2\text{O}_7)\text{K}_2\text{Cr}_2\text{O}_7) \cdot V(\text{K}_2\text{Cr}_2\text{O}_7) \cdot M(f_{\text{экв.}}(\text{Fe})\text{Fe})}{1000} \cdot \frac{V_{\text{м.к.}}}{V^a(\text{Fe}^{2+})} \cdot \frac{100}{m(\text{руды})}.$$

$$F_{\text{экв.}}(\text{K}_2\text{Cr}_2\text{O}_7) = 1/6; \quad f_{\text{экв.}}(\text{Fe}^{2+}) = 1; \quad M(\text{Fe}^{2+}) = 55,85 \text{ г/моль.}$$

$$w(\%)(\text{Fe}) = \frac{0,04959 \cdot 6,61 \cdot 55,85}{1000} \cdot \frac{200,0}{9,95} \cdot \frac{100}{0,6601} = 55,75 \text{ .}$$

4. Каково относительное стандартное отклонение результата анализа?

Относительное стандартное отклонение результата анализа рассчитываем по формуле:

$$\frac{s_y}{y} = \sqrt{\sum_{i=1}^n \left(\frac{s_{x_i}}{x_i} \right)^2};$$

$$\frac{s_{w(\%)(\text{Fe})}}{w(\%)(\text{Fe})} = \sqrt{\left(\frac{s_{c(1/6\text{K}_2\text{Cr}_2\text{O}_7)}}{c(1/6\text{K}_2\text{Cr}_2\text{O}_7)} \right)^2 + \left(\frac{s_{\bar{V}(\text{K}_2\text{Cr}_2\text{O}_7)}}{\bar{V}(\text{K}_2\text{Cr}_2\text{O}_7)} \right)^2 + \left(\frac{s_{V_{\text{м.к.}}}}{V_{\text{м.к.}}} \right)^2 + \left(\frac{s_{m(\text{руды})}}{m(\text{руды})} \right)^2}$$

$$\frac{s_{w(\%)(\text{Fe})}}{w(\%)(\text{Fe})} = \sqrt{\left(\frac{9 \cdot 10^{-5}}{0,04959} \right)^2 + \left(\frac{2,12 \cdot 10^{-2}}{6,614} \right)^2 + \left(\frac{2 \cdot 10^{-1}}{200,0} \right)^2 + \left(\frac{4 \cdot 10^{-4}}{0,6601} \right)^2} =$$

⁶⁾ Стандартное отклонение с равной вероятностью может быть как положительным, так и отрицательным, поскольку оно рассчитывается как квадратный корень из дисперсии $s =$

$\sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$, поэтому значение стандартного отклонения записываем со знаком ±.

$$= \sqrt{3,29 \cdot 10^{-6} + 10,27 \cdot 10^{-6} + 1 \cdot 10^{-6} + 0,367 \cdot 10^{-6}} = 3,86 \cdot 10^{-3}.$$

5. Каково абсолютное стандартное отклонение?

$$s_{w(\%)(\text{Fe})} = 3,86 \cdot 10^{-3} \cdot 55,75 = 0,22 (\%).$$

6. Каково значение результата анализа?

$$w(\%)(\text{Fe}) \pm s_{w(\%)(\text{Fe})} = 55,75 \pm 0,22.$$

Число 0,22 показывает, что случайная погрешность в определении массовой доли железа в руде присутствует уже в первом десятичном знаке после запятой, поэтому результаты округляем таким образом: $w(\%)(\text{Fe}) = 55,8 \pm 0,2$.

Доверительный интервал косвенно измеряемой величины (в нашем примере $w(\%)(\text{Fe})$) можно получить из выражения: $\Delta \bar{y} = t_{p,f} \cdot \frac{s_y}{\sqrt{n}}$, где n – число аргументов x , однако при обработке результатов косвенных измерений обычно ограничиваются только расчетом дисперсии s_y^2 или стандартного отклонения s_y .

Ответ: Абсолютное стандартное отклонение результата анализа равно 0,2 %; массовая доля (%) железа в руде составляет $55,8 \pm 0,2$.

Пример 2.9 - Растворимость (S) хлорида серебра в воде равна $(1,33 \pm 0,004) \cdot 10^{-5}$ моль/дм³. Какова абсолютная случайная погрешность вычисления произведения растворимости? (В скобках указано абсолютное стандартное отклонение растворимости).

Решение

1. Каково значение произведения растворимости хлорида серебра?

$$\text{ПР}(\text{AgCl}) = [\text{Ag}^+][\text{Cl}^-] = S^2(\text{AgCl});$$

$$\text{ПР}(\text{AgCl}) = (1,33 \cdot 10^{-5})^2 = 1,77 \cdot 10^{-10}.$$

2. Каково относительное стандартное отклонение результата?

$$\left| \frac{s_y}{y} \right| = \left| n \cdot \frac{s_x}{x} \right|; \quad \left| \frac{s_{\text{ПР}(\text{AgCl})}}{\text{ПР}(\text{AgCl})} \right| = \left| 2 \cdot \frac{s_{S(\text{AgCl})}}{S(\text{AgCl})} \right|.$$

$$\left| \frac{s_{\text{ПР(AgCl)}}}{\text{ПР(AgCl)}} \right| = \frac{2 \cdot 0,004 \cdot 10^{-5}}{1,33 \cdot 10^{-5}} = 6,0 \cdot 10^{-3}.$$

3. Каково абсолютное стандартное отклонение результата?

$$s_{\text{ПР(AgCl)}} = 6,0 \cdot 10^{-3} \cdot 1,77 \cdot 10^{-10} = 0,01 \cdot 10^{-10}.$$

4. Каково значение результата?

$$\text{ПР(AgCl)} \pm s_{\text{ПР(AgCl)}} = 1,77 \cdot 10^{-10} \pm 0,01 \cdot 10^{-10}.$$

Ответ: Абсолютная случайная погрешность результата определения ПР(AgCl) равняется $0,01 \cdot 10^{-10}$, а $\text{ПР(AgCl)} = (1,77 \pm 0,01) \cdot 10^{-10}$.

2.7 Статистическая оценка предела обнаружения

Одной из основных проблем аналитической химии является определение «следовых» количеств примесей, которые могут существенно влиять на свойства исследуемого материала или обладать токсическим действием. При определении очень низких концентраций мешающее влияние оказывают примеси в реактивах. Хотя концентрация определяемого микрокомпонента (x) в «холостом» опыте должна быть равна нулю, обычно получают отличающееся от нуля значение аналитического сигнала (фона) (y_0), например, поглощение света «холостым» раствором в фотометрии. Если исследовать много (n_0) «холостых» проб (где $x \approx 0$), то получают соответствующее количество аналитических сигналов, которые рассеиваются вблизи среднего значения аналитического сигнала «холостого» опыта (\bar{y}_0) со стандартным отклонением аналитического сигнала s_0 . Минимальное значение аналитического сигнала исследуемой пробы (\bar{y}_{\min}) (среднее из n параллельных определений), значимо отличающееся от среднего результата «холостого» опыта (\bar{y}_0), должно выходить за пределы границ доверительного интервала аналитического сигнала фона:

$$\bar{y}_{min} > \bar{y}_o + \Delta \bar{y}_o \quad \text{или} \quad \bar{y}_{min} > \bar{y}_o + t_{P,f} \cdot \frac{s_{y_o}}{\sqrt{n_o}} \quad \text{или} \quad \bar{y}_{min} > \bar{y}_o + t_{P,f} \cdot s_{y_o}$$

где s_{y_o} – стандартное отклонение для среднего значения аналитического сигнала «холостого» опыта.

Для получения результата необходимо пересчитать содержание определяемого компонента \bar{x}_{min} на значение аналитического сигнала \bar{y}_{min} . Если между результатом анализа (\bar{x}_i) и аналитическим сигналом (\bar{y}_i) наблюдается линейная зависимость (раздел 3.1)

$$y = a + bx,$$

где $a = \bar{y}_o$ (при $x = 0$ - среднее значение аналитического сигнала «холостого» опыта);

b – инструментальный коэффициент чувствительности (S) методики анализа ($\text{tg}\alpha$, если α - угол наклона прямой),

то получим

$$\bar{x}_{min} > \frac{\bar{y}_{min} - \bar{y}_o}{S} = \frac{t_{P,f} \cdot s_{y_o}}{S \cdot \sqrt{n_o}}.$$

Наименьшее содержание определяемого компонента ($m_{min,P}$ – минимальная масса или $c_{min,P}$ – минимальная концентрация), которое можно обнаружить с заданной доверительной вероятностью, называется пределом обнаружения. Предел обнаружения является обязательной метрологической характеристикой методик определения микроконцентраций элементов.

Расчетные значения предела обнаружения для одного и того же элемента при одной и той же методике могут отличаться на порядок и более. Это обусловлено отсутствием единого мнения о значении доверительной вероятности P , которую необходимо использовать при вычислении $m_{min,P}$ (или $c_{min,P}$), и о методике расчета стандартного отклонения аналитического сигнала «холостого» опыта.

Если проводить большое число параллельных определений аналитического сигнала фона ($n \geq 20$), то коэффициент Стьюдента ($t_{P, f}$) при доверительной вероятности $P = 0,999$ принимает значение, близкое к 3, а при $P = 0,95$ – значение, близкое к 2. Поэтому формула для определения предела обнаружения обычно имеет вид:

$$\bar{x}_{min} = \frac{3s_{y_0}}{S} \quad (\text{при } P = 0,999) \text{ или}$$

$$\bar{x}_{min} = \frac{2s_{y_0}}{S} \quad (\text{при } P = 0,95) .$$

Существуют другие способы вычисления предела обнаружения, но формулы, приведенные выше, используют чаще всего.

Пример 2.10 «Следы» железа в чистом алюминии определяли фотометрически с сульфосалициловой кислотой в кювете с толщиной слоя $l = 0,01$ м (молярный коэффициент светопоглощения $\varepsilon = 5,6 \cdot 10^2 \text{ м}^2/\text{моль}$). Среднее значение оптической плотности «холостого» раствора равняется 0,08, стандартное отклонение для среднего значения аналитического сигнала «холостого» опыта $s_{y_0} = 0,02$ (при $n_0 = 20$). Определить предел обнаружения.

Решение

1. Каково значение коэффициента чувствительности (S)?

Из закона Бугера-Ламберта-Бера следует, что $A = \varepsilon \cdot c \cdot l$.

$$S = \frac{\Delta A}{\Delta c} = \varepsilon \cdot l ; \quad S = 5,6 \cdot 10^2 \cdot 0,01 = 5,6 \text{ (м}^3/\text{моль)}.$$

2. Каково значение предела обнаружения?

$$\bar{c}_{min} = \frac{3s_{y_0}}{S} \quad (\text{при } P = 0,999);$$

$$\bar{c}_{min} = \frac{3 \cdot 0,02}{5,6} = 1,07 \cdot 10^{-2} \text{ (моль/м}^3\text{)} \quad \text{или} \quad 1,07 \cdot 10^{-5} \text{ моль/дм}^3 .$$

Ответ: Предел обнаружения $\bar{c}_{min} = 1,07 \cdot 10^{-5} \text{ моль/дм}^3$.

Предел обнаружения можно понизить, уменьшая стандартное отклонение для среднего значения аналитического сигнала „холостого” опыта за счет увеличения числа измерений аналитического сигнала фона или повышения инструментального коэффициента чувствительности S .

2.8 Контрольные вопросы

1. Дать определение генеральной и выборочной совокупности (выборки).
2. Дать понятие воспроизводимости анализа; критерии ее оценки.
3. Привести определение дисперсии и стандартного отклонения.
4. Дать понятие числа степеней свободы (f).
5. Описать ход вычисления дисперсии и стандартного отклонения для генеральной и выборочной совокупностей.
6. Показать связь между дисперсией единичных результатов и дисперсией среднего.
7. Дать определение относительного стандартного отклонения, коэффициента вариации.
8. Описать ход построения гистограммы, проверки нормальности распределения результатов анализа.
9. Дать определения понятиям доверительный интервал, границы доверительного интервала и интервальное значение измеряемой величины.
10. Описать ход расчета доверительного интервала для единичного результата и среднего при заданной доверительной вероятности.
11. Указать содержание нормированных коэффициентов U и t .
12. Привести статистические критерии обнаружения промахов.
13. Описать способ обнаружения систематической погрешности по действительным значениям и статистическим оценкам выборки.

14. Привести содержание понятия нуль-гипотеза; роль нуль-гипотезы в проверке наличия систематической погрешности.

15. Привести алгоритм статистической обработки результатов химического анализа.

16. Описать способ статистической проверки значимости расхождения дисперсий двух методов.

17. Дать расчет средневзвешенной дисперсии двух выборочных совокупностей.

18. Привести известные законы распределения случайных величин.

19. Привести правила вычисления суммарной случайной погрешности косвенных измерений.

20. Дать понятие предела обнаружения; формула для его вычисления.

3 Определение содержания компонента по градуировочному графику

В большинстве физико-химических и физических методов анализа содержание компонента находят с помощью градуировочных графиков. В системе координат (xOy), где x – содержание компонента, y – аналитический сигнал, строят график с использованием стандартных образцов с точно известным содержанием определяемого компонента (рисунок 6). После измерения величины аналитического сигнала в исследуемой пробе по нему находят содержание определяемого компонента.

Градуировочный график имеет, как правило, линейный характер в довольно широком диапазоне содержания определяемого компонента. При его построении может наблюдаться значительное рассеяние результатов измерений, в особенности при работе с малыми концентрациями веществ. Это затрудняет определение правильного хода градуировочного графика.

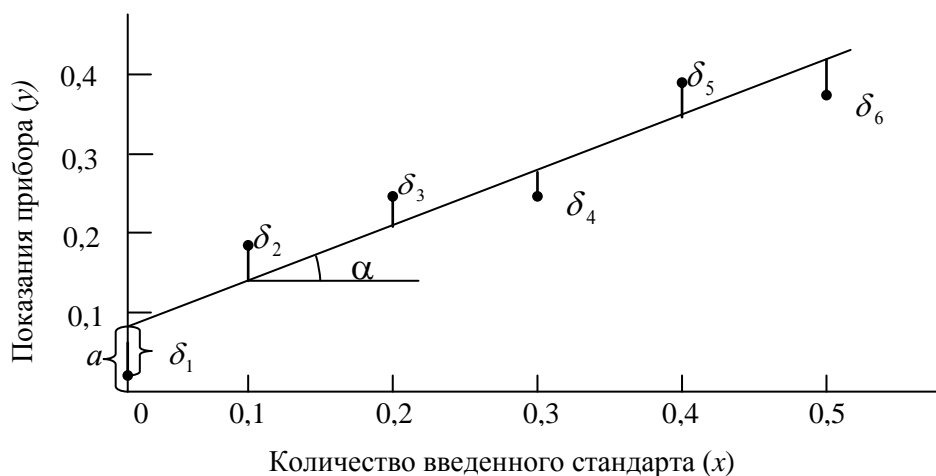


Рисунок 6 - Типичный градуировочный график

Если величина y линейно зависит от величины x , то для правильного проведения прямой, уравнение которой выражается зависимостью

$$Y = a + bx, \quad (17)$$

необходимо использовать линию тренда с выводением параметра R^2 . Разность между измеренными величинами y_i и вычисленными значениями Y_i (δ_i) по уравнению (17) должна быть наименьшей. Оптимальную прямую можно найти графически или аналитически – методом наименьших квадратов.

При графическом построении точки выбирают программную среду Excel, удобную для построения графиков и выполнения вычислений.

Постоянный параметр a в уравнении (17) находят как отрезок на оси ординат при $x = 0$, а величина параметра b определяется как тангенс угла наклона прямой ($b = \operatorname{tg} \alpha$) к положительному направлению оси Ox . Отрезок a на оси Oy отвечает значению «холостого» опыта, а b коэффициент чувствительности (S) методики анализа или инструментальный коэффициент чувствительности ($S = \operatorname{tg} \alpha$). Содержание определяемого компонента (x) находят, пользуясь градуировочным графиком или уравнением (17), параметры a и b которого уже известны.

3.1 Подбор параметров a и b методом наименьших квадратов для построения градуировочного графика

Для получения градуировочной прямой используют стандартный образец или набор образцов. Задача нахождения формулы разбивается на два этапа. Сначала устанавливают характер зависимости $y = f(x)$, определяют, является ли зависимость $y = f(x)$ линейной, квадратичной, показательной, степенной и т.п. На втором этапе необходимо подбирают параметры функции так, чтобы функция $f(x)$ наилучшим образом соответствовала экспериментальным данным.

При поиске уравнений исходят из предположения, что совокупность экспериментальных точек описывается одним из известных типов уравнений.

Прежде чем определять значения коэффициентов в выбранном уравнении, необходимо проверить возможность его использования методом выравнивания, который заключается в преобразовании функции $y=f(x)$ в линейную функцию. Это достигается путем замены переменных x и y новыми переменными, которые подбираются таким образом, чтобы получить уравнение прямой. Так, замена величины $1/x$ на новую переменную z в уравнении $y = a + b/x$ приводит его к виду $Y = a + bz$. Аналогично в уравнении $y = a + bx^n$ замена x^n на новую переменную z преобразует степенную зависимость в линейную $Y = a + bz$. При логарифмировании показательная зависимость $y = ax^n$ превращается в линейную $\lg y = \lg a + n \cdot \lg x$. Любая зависимость, которая имеет вид степенного двучлена $y = ax^m + bx^n$, введением двух новых переменных $t = y/x^m$ и $V = x^{n-m}$ может быть сведена к линейной $t = a + bV$.

Вычислив по известным значениям x и y значения новых переменных, наносят их на координатную плоскость. Если нанесенные точки располагаются вблизи прямой линии, то выбранное эмпирическое уравнение подходит для характеристики зависимости $y=f(x)$.

При нахождении неизвестных параметров a и b функции чаще всего пользуются так называемым принципом наименьших квадратов. Он состоит в том, что из ряда формул вида $y = f(x, a, b, c, \dots)$, где a, b, c, \dots – неизвестные параметры, лучше всего соответствует экспериментальным данным та, для которой сумма квадратов отклонений (SQ) ординат экспериментальных точек от ординат соответствующих точек аппроксимирующей прямой минимальна:

$$SQ = \sum_{i=1}^n \delta_{y_i}^2 = \sum_{i=1}^n (f(x_i, a, b, c, \dots) - y_i)^2 \rightarrow \min. \quad (18)$$

Функция $SQ = SQ(a, b, c, \dots)$ является функцией переменных a, b, c, \dots, x_i и y_i – числа, найденные экспериментально. Необходимое условие экстремума функции $SQ(a, b, c, \dots)$ – равенство нулю частных производных функции по каждому из параметров. Значения a, b, c, \dots из системы уравнений, которая

называется нормальной, подбираются так, чтобы функция $SQ(a, b, c, \dots)$ достигала минимума.

Нахождение параметров функции $y = f(x, a, b, c, \dots)$ по этому принципу называется методом наименьших квадратов после того, как вид функции $y = f(x, a, b, c, \dots)$ определен.

После подбора уравнения методом наименьших квадратов определяют его коэффициенты и получают эмпирическую формулу, которая представляет экспериментальные данные в аналитическом виде. Для определения параметров эмпирической формулы необходимо иметь таблицу экспериментальных данных зависимостей между x и y .

Метод наименьших квадратов является основным методом раздела математики, который получил название регрессионный анализ. Его задачей является установление вида самой функции, то есть значений параметров уравнения по известной совокупности значений x_i, y_i .

Рассмотрим линейную зависимость между содержанием определяемого компонента и аналитическим сигналом (17), градуировочный график которой используется для определения содержания компонента. Задача сводится к нахождению параметров a и b уравнения связи двух переменных (парной линейной регрессии).

Истинные значения этих параметров найти нельзя при ограниченной выборке экспериментальных данных; $n = 5-7$ (5-7 экспериментальных точек). Рассчитать параметры a и b можно методом наименьших квадратов. При этом должны выполняться следующие условия:

1. Значения x задаются без погрешностей или их погрешности намного меньше погрешностей измерения значений y , то есть погрешностями измерений значений x можно пренебречь.

2. Измеренные значения y_i и отклонения $\delta_{y_i} = y_i - Y_i$, (разность между измеренным y_i и вычисленным Y_i значением) аналитического сигнала при одном и

том же содержании определяемого компонента (x_i), должны подчиняться закону нормального распределения.

В этом случае в качестве наилучших значений параметров a и b выбирают такие, для которых сумма квадратов отклонений $\delta_{y_i}^2$ минимальна. Аналитически задача метода наименьших квадратов для линейной зависимости формулируется следующим образом:

$$\sum_{i=1}^n \delta_{y_i}^2 = \sum_{i=1}^n (Y_i - y_i)^2 = \sum [(a + bx_i) - y_i]^2 \rightarrow \min . \quad (19)$$

В дальнейшем с целью упрощения символ суммы $\sum_{i=1}^n$ будет обозначен как \sum .

Учитывая обозначение $\sum_{i=1}^n \delta_{y_i}^2$ как SQ , получим:

$$SQ = \sum (a + bx_i - y_i)^2 , \quad (20)$$

функция $SQ = SQ(a, b)$ является функцией двух переменных a и b , а x_i и y_i – постоянные числа, найденные экспериментально. Минимум функции определяется по правилам дифференциального исчисления функции двух переменных из системы:

$$\begin{cases} SQ'_a = \frac{\partial SQ}{\partial a} = 0 \\ SQ'_b = \frac{\partial SQ}{\partial b} = 0 \end{cases} , \quad (21)$$

которая выражает необходимые условия экстремума (SQ'_a и SQ'_b – частные производные первого порядка по переменным a и b , соответственно).

Найдем SQ'_a и SQ'_b для функции $SQ = \sum (a + bx_i - y_i)^2$ и запишем соответствующую систему:

$$\begin{cases} SQ'_a = 2 \sum (a + bx_i - y_i) = 2(an + b \sum x_i - \sum y_i) \\ SQ'_b = 2 \sum (a + bx_i - y_i)x_i = 2(a \sum x_i + b \sum x_i^2 - \sum x_i y_i). \end{cases} \quad (22)$$

$$\begin{cases} 2(an + b \sum x_i - \sum y_i) = 0 \\ 2(a \sum x_i + b \sum x_i^2 - \sum x_i y_i) = 0 \end{cases} \quad (23)$$

или

$$\begin{cases} b \sum x_i^2 + a \sum x_i = \sum x_i y_i \\ b \sum x_i + an = \sum y_i. \end{cases} \quad (24)$$

Система (23) называется системой нормальных уравнений.

Для нахождения параметров a и b можно использовать метод Крамера или Гаусса. Эта система имеет единственное решение, поскольку ее определитель (Δ) не равен нулю:

$$\Delta = \begin{vmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{vmatrix} = n \sum x_i^2 - (\sum x_i)^2 \neq 0 \quad (25)$$

(точнее, $\Delta > 0$, что можно доказать методом математической индукции при $n \geq 2$).

Из уравнений (25), $\Delta = n \sum x_i^2 - (\sum x_i)^2$, отсюда найдем Δ_1 и Δ_2 и по формулам

Крамера найдем $a = \frac{\Delta_2}{\Delta}$ и $b = \frac{\Delta_1}{\Delta}$.

$$\Delta_1 = \begin{vmatrix} \sum x_i y_i & \sum x_i \\ \sum y_i & n \end{vmatrix} = n \sum x_i y_i - \sum x_i \cdot \sum y_i \quad (26)$$

$$\Delta_2 = \begin{vmatrix} \sum x_i^2 & \sum x_i y_i \\ \sum x_i & \sum y_i \end{vmatrix} = \sum x_i^2 \cdot \sum y_i - \sum x_i \cdot \sum x_i y_i. \quad (27)$$

Формулы для нахождения a и b имеют вид:

$$a = \frac{\sum x_i^2 \cdot \sum y_i - \sum x_i \cdot \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2}, \quad (28)$$

$$b = \frac{n \sum x_i y_i - \sum x_i \cdot \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}. \quad (29)$$

Примечание - Используя достаточное условие экстремума функции двух переменных, можно проверить, что функция $SQ = SQ(a, b)$ имеет единственную точку минимума, определенную из системы нормальных уравнений, и в этой точке функция имеет не только локальный минимум, но и наименьшее значение (глобальный минимум).

Определив параметры a и b , получаем эмпирическое уравнение $Y = a + bx$, которое используют для нахождения содержания определяемого компонента. Для этого строят градуировочный график, подставляя исходные значения x_i и рассчитанные значения Y_i .

Применение метода Гаусса для решения системы линейных уравнений, суть которого состоит в последовательном исключении неизвестных переменных. Чаще всего кривые на малых участках аппроксимируют полиномом второго порядка:

$$V = a + bT + cT^2 \quad (30)$$

и определяют коэффициенты этого уравнения a, b, c .

Сначала исключается коэффициент a из всех уравнений вычитая из второго и третьего уравнение (30). Получают систему двух уравнений с двумя неизвестными b и c , из которой исключают b , затем находят c .

Зная c , восстанавливают b и a из первого уравнения:

$$c = \frac{(T_3 - T_1)(V_2 - V_1) - (T_2 - T_1)(V_3 - V_1)}{(T_3^2 - T_1^2)(T_3 - T_1) - (T_3^2 - T_1^2)(T_2 - T_1)}$$

$$b = \frac{(V_3 - V_1) - c(T_2^2 - T_1^2)}{(T_2 - T_1)}$$

$$a = V_1 - bT_1 - cT_1^2 \quad . \quad (31)$$

По градуировочному графику определяют метрологические характеристики параметров:

- дисперсию линии регрессии,
- дисперсии параметров a и b ,
- значимость параметра a ,
- доверительные интервалы и интервальные значения параметров a и b .

Дисперсия линии регрессии s_0^2 характеризует рассеяние экспериментальных точек (x_i, y_i) относительно прямой. Ее вычисляют по формуле:

$$s_0^2 = \frac{\sum (y_i - Y_i)^2}{n - 2} \quad (32)$$

при числе степеней свободы $f = n - 2$.

Поскольку параметры a и b являются случайными величинами, то необходима оценка их доверительных интервалов. Оценка доверительных

интервалов параметров a и b проводится на основе значений *стандартных отклонений* s_a и s_b , которые связаны с дисперсиями параметров a и b следующим образом:

$$s_a = \sqrt{s_a^2}; \quad s_b = \sqrt{s_b^2}. \quad (33)$$

Дисперсии параметров a и b (s_a^2 и s_b^2) можно определить, зная

$$s_a^2 = \frac{s_0^2 \cdot \sum x_i^2}{n \sum (x_i - \bar{x})^2} = \frac{s_0^2 \cdot \sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2}; \quad s_b^2 = \frac{s_0^2}{\sum (x_i - \bar{x})^2} = \frac{ns_0^2}{n \sum x_i^2 - (\sum x_i)^2}. \quad (34)$$

Если значение параметра a (отрезок, отсекаемый на оси ординат) мало отличается от нуля, проверяют его значимость, определяют, обусловлено ли неравенство $a \neq 0$ недостаточной воспроизводимостью или систематической погрешностью.

Простейший метод оценки значимости a – метод Юдена, в котором вычисляют значение коэффициента Стьюдента:

$$t_{a \text{ расч.}} = \frac{|a|}{s_a}, \quad (35)$$

где s_a – стандартное отклонение параметра a .

Сравнивают $t_{a \text{ расч.}}$ с $t_{\text{табл.}}$ (P, f) (табличным значением коэффициента Стьюдента). Значение $t_{\text{табл.}}$ находят по таблице приложения Б при доверительной вероятности $P = 0,95$ и числе степеней свободы $f = n - 2$.

Если $t_{a \text{ расч.}} > t_{\text{табл.}}$, то a – значимо, величиной a нельзя пренебречь. В этом случае считают, что появление величины a за счет одних случайных погрешностей маловероятно, метод содержит систематическую погрешность, которая равна a .

Если $t_{a \text{ расч.}} < t_{\text{табл.}}$, то a – незначимо, величиной a в уравнении (16) можно пренебречь и заново найти угловой коэффициент b' прямой

$$Y' = b'x, \quad (36)$$

проходящей через начало координат, значение Y'_i и дисперсию параметра b' по формулам:

$$b' = \frac{\sum x_i y_i}{\sum x_i^2}; \quad (37)$$

$$s_{b'}^2 = \frac{s_0^2}{\sum x_i^2}. \quad (38)$$

Доверительные интервалы параметров a и b рассчитывают по формулам:

$$\pm \Delta a = t_{P, f} \cdot s_a; \quad (39)$$

$$\pm \Delta b = t_{P, f} \cdot s_b. \quad (40)$$

Величину $t_{P, f}$ находят по таблице приложения 2 при $P = 0,95$ и $f = n - 2$. Если $a = 0$, то при определении доверительного интервала параметра b значение $t_{P, f}$ находят при $f = n - 1$.

Интервальные значения параметров a и b находят по формулам:

$$a \pm \Delta a = a \pm t_{P, f} \cdot s_a, \quad (41)$$

$$b \pm \Delta b = b \pm t_{P, f} \cdot s_b. \quad (42)$$

При вычислениях по формулам 30 - 42 все расчеты необходимо выполнять с бóльшим количеством знаков после запятой, в особенности при нахождении разности $(y_i - Y_i)$, так как это есть разность между близкими по величине значениями. Округления в ходе промежуточных расчетов приводят к большим погрешностям вычислений. Количество значащих цифр, сохраняемых в конечном результате при указанных значениях параметров ***a*** и ***b***, определяется величиной доверительного интервала параметров ***a*** и ***b***, то есть, зная Δa и Δb , определяют число необходимых знаков после запятой для значений ***a*** и ***b***.

3.2 Определение содержания компонента с помощью уравнения линейной регрессии и градуировочного графика

Зная метрологические характеристики параметров ***a*** и ***b***, можно перейти к расчету содержания определяемого компонента и его метрологических характеристик. Определяемая концентрация компонента должна обязательно находиться в интервале концентраций, для которых рассчитано уравнение. При этом условии измеряют значение аналитического сигнала исследуемого раствора ($y_{\text{контр.}}$) и по уравнению (43), полученному из уравнения (30), находят содержание определяемого компонента ($x_{\text{контр.}}$) в образце:

$$x_{\text{контр.}} = \frac{y_{\text{контр.}} - a}{b}. \quad (43)$$

Значение аналитического сигнала $y_{\text{контр.}}$ является случайной величиной и для оценки воспроизводимости результатов проводят m параллельных измерений в контрольной пробе ($y_{\text{контр. 1}}, y_{\text{контр. 2}}, \dots, y_{\text{контр. } m}$) или измерения аналитических сигналов для m параллельно анализируемых проб ($m \geq 3$). Тогда

$$\bar{y}_{\text{контр.}} = \frac{\sum_{i=1}^m y_{i\text{контр.}}}{m}, \quad (44)$$

$$\bar{x}_{\text{контр.}} = \frac{(\bar{y}_{\text{контр.}} - a)}{b}. \quad (45)$$

Для этого сначала находим стандартное отклонение (s) результата анализа.

Стандартное отклонение (s) для ($x_{\text{контр.}}$) рассчитывается по формуле (9):

$$s_{x_{\text{контр.}}} = \frac{s_0}{b} \left[\frac{1}{m} + \frac{1}{n} + \frac{n(\bar{y}_{\text{контр.}} - \bar{y})^2}{b^2 [n \sum x_i^2 - (\sum x_i)^2]} \right]^{1/2}. \quad (46)$$

Формула (44) может быть преобразована к виду:

$$s_{x_{\text{контр.}}} = \frac{s_0}{b} \left[\frac{1}{m} + \frac{1}{n} + \frac{(\bar{y}_{\text{контр.}} - \bar{y})^2}{b^2 \sum (x_i - \bar{x})^2} \right]^{1/2} \quad (47)$$

или

$$s_{x_{\text{контр.}}} = \frac{s_0}{b} \left[\frac{1}{m} + \frac{1}{n} + \frac{(\bar{y}_{\text{контр.}} - \bar{y})^2}{b^2 (\sum x_i^2 - n \bar{x}^2)} \right]^{1/2}, \quad (48)$$

где n – число точек при построении градуировочного графика;

m – число параллельных измерений аналитического сигнала в контрольном (исследуемом) растворе;

x_i – значения содержания определяемого компонента в стандартном образце;

\bar{x} – среднее из n значений содержания определяемого компонента в стандартных образцах, взятых для построения градуировочного графика;

\bar{y} – среднее из n значений аналитических сигналов стандартных образцов.

Доверительный интервал для среднего значения результата анализа ($\Delta \bar{x}_{\text{контр.}}$) рассчитывают по формуле:

$$\Delta \bar{x}_{\text{контр.}} = t_{P, f} \cdot s_{x_{\text{контр.}}} , \quad (49)$$

где $s_{x_{\text{контр.}}}$ вычисляют по одной из формул (46), (47), (48);

$t_{P, f}$ – значение коэффициента Стьюдента, его находят из таблицы приложения 2 при числе степеней свободы f , равном $n - 2$, и, как правило, при доверительной вероятности P , равной 0,95.

Интервальные значения содержания определяемого компонента в образце оценивают по общей формуле:

$$\bar{x}_{\text{контр.}} \pm \Delta \bar{x}_{\text{контр.}} . \quad (50)$$

Используя приведенные выше сведения, решим конкретную задачу.

Пример. Расчет содержания определяемого компонента с помощью градуировочного графика

Рассчитать среднее и интервальные значения массы Zn^{2+} -ионов в контрольной пробе, а также относительную погрешность по данным фотоколориметрических определений растворов, содержащих Zn^{2+} -ионы, представленных в таблице 6.

Таблица 6 - Данные фотометрического определения Zn^{2+} - ионов

$m(Zn^{2+}), \text{мкг } (x_i)$	0,00	0,10	0,20	0,30	0,40	0,50	$x_{\text{контр.}}$
$A \ (y_i)$	0,020	0,120	0,170	0,230	0,290	0,330	$y_{\text{контр.1}} = 0,255$ $y_{\text{контр.2}} = 0,260$ $y_{\text{контр.3}} = 0,265$

Решение

При расчете параметров градуировочного графика рекомендуется располагать материал так, как показано в таблице 7.

Таблица 7 - Исходные данные и вспомогательные расчеты при оценке параметров градуировочного графика определения Zn^{2+} -ионов

y_i	x_i	x_i^2	$x_i y_i$	Y_i
0,020	0,00	0,000	0,0	0,042
0,120	0,10	0,010	0,012	0,103
0,170	0,20	0,040	0,034	0,163
0,230	0,30	0,090	0,069	0,224
0,290	0,40	0,160	0,116	0,284
0,330	0,50	0,250	0,165	0,345
$\sum_1^n y_i = 1,160$	$\sum_1^n x_i = 1,50$	$\sum_1^n x_i^2 = 0,550$	$\sum_1^n x_i y_i = 0,396$	

$$a = \frac{\sum_1^n x_i^2 \cdot \sum_1^n y_i - \sum_1^n x_i \cdot \sum_1^n x_i y_i}{n \cdot \sum_1^n x_i^2 - \left(\sum_1^n x_i \right)^2};$$

$$a = \frac{0,550 \cdot 1,160 - 1,50 \cdot 0,396}{6 \cdot 0,550 - (1,50)^2} = 0,0419.$$

$$b = \frac{n \cdot \sum_1^n x_i y_i - \sum_1^n x_i \cdot \sum_1^n y_i}{n \cdot \sum_1^n x_i^2 - \left(\sum_1^n x_i \right)^2};$$

$$b = \frac{6 \cdot 0,396 - 1,50 \cdot 1,160}{6 \cdot 0,550 - (1,50)^2} = 0,606.$$

Получаем первоначальное уравнение градуировочного графика:

$$Y = 0,0419 + 0,606x,$$

параметры которого обрабатываем статистически:

1. Рассчитываем дисперсию линии регрессии s_0^2 по формуле:

$$s_0^2 = \frac{\sum (y_i - Y_i)^2}{n-2}.$$

$$s_0^2 = \frac{0,022^2 + 0,017^2 + 0,007^2 + 0,006^2 + 0,006^2 + 0,015^2}{4} = 2,8 \cdot 10^{-4}.$$

2. Рассчитываем дисперсии параметров a и b (s_a^2 и s_b^2) по формулам:

$$s_a^2 = \frac{s_0^2 \cdot \sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2}; \quad s_a^2 = \frac{2,8 \cdot 10^{-4} \cdot 0,550}{6 \cdot 0,550 - (1,50)^2} = 1,5 \cdot 10^{-4}.$$

$$s_b^2 = \frac{n s_0^2}{n \sum x_i^2 - (\sum x_i)^2}; \quad s_b^2 = \frac{6 \cdot 2,8 \cdot 10^{-4}}{6 \cdot 0,550 - (1,50)^2} = 1,6 \cdot 10^{-3}.$$

3. Проверяем значимость параметра a

$$t_{a \text{ расч.}} = \frac{|a|}{s_a} = \frac{0,0419}{1,2 \cdot 10^{-2}} = 3,48; \quad t_{a \text{ табл.}} (P=0,95; f=4) = 2,78$$

$t_{a \text{ расч.}} > t_{a \text{ табл.}}$, следовательно a значимо.

4. Рассчитываем доверительный интервал параметров a и b

$$\pm \Delta a = t_{P,f} \cdot s_a = 2,78 \cdot 1,2 \cdot 10^{-2} = 0,033;$$

$$\pm \Delta b = t_{P,f} \cdot s_b = 2,78 \cdot 4,0 \cdot 10^{-2} = 0,111.$$

Учитывая, что погрешность в определении параметра a будет во втором десятичном знаке после запятой, а параметра b – в первом десятичном знаке после запятой, значение параметра a (0,0419) округляем до 0,04, а значение параметра b (0,606) округляем до 0,6.

Таким образом, эмпирическое уравнение прямой градуировочного графика фотометрического определения Zn^{2+} -ионов можно записать в виде:

$$Y = 0,04 + 0,6x \quad \text{или} \quad A = 0,04 + 0,6 m(Zn^{2+}).$$

Экспериментальные и рассчитанные значения оптической плотности раствора при соответствующих значениях массы Zn^{2+} -ионов представлены на рисунке 7 графиком зависимости $Y = 0,04 + 0,6x$.

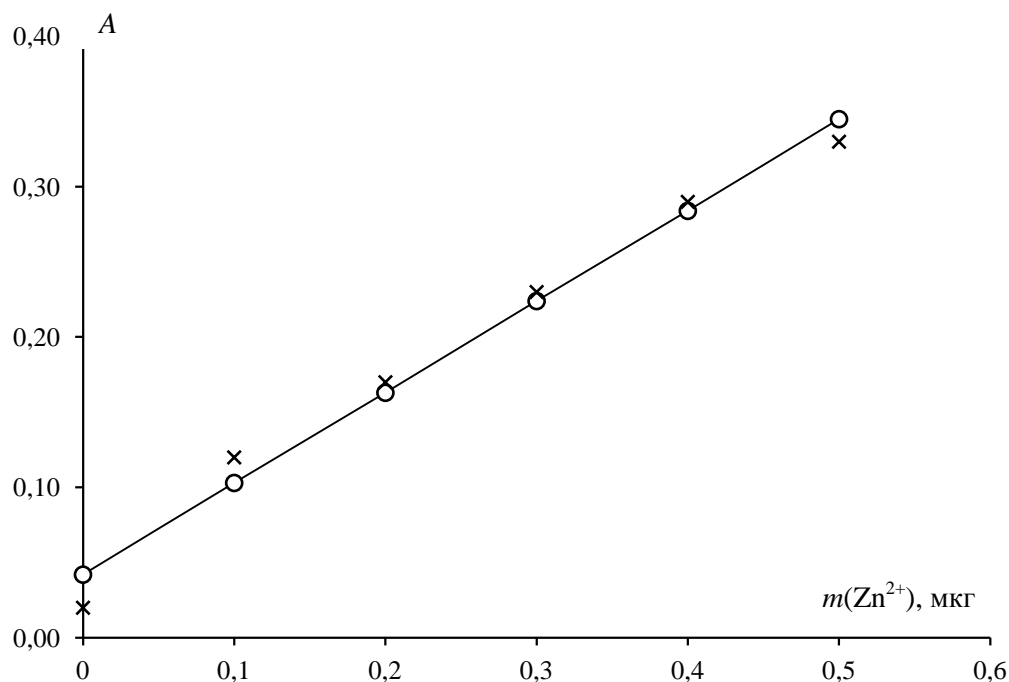


Рисунок 7 - Градуировочный график фотометрического определения Zn^{2+} -ионов (\times – экспериментальные; \circ – рассчитанные точки)

Расчет коэффициентов a и b уравнения градуировочного графика $Y = a + bx$ (линии регрессии) методом наименьших квадратов, расчет дисперсии линии регрессии и дисперсии параметров a и b , проверка значимости параметра a обычно выполняется с помощью компьютера.

Нахождение содержания определяемого компонента. Расчет содержания определяемого компонента x можно провести по графику или по уравнению (44) или (45). Рассмотрим расчет неизвестной концентрации вещества по уравнению (45):

$$\bar{y}_{\text{контр.}} = \frac{\sum_{i=1}^m y_{i_{\text{контр.}}}}{m}; \quad \bar{y}_{\text{контр.}} = \frac{0,255 + 0,260 + 0,265}{3} = 0,260.$$

$$\bar{x}_{\text{контр.}} = \frac{(\bar{y}_{\text{контр.}} - a)}{b}; \quad \bar{x}_{\text{контр.}} = \frac{0,260 - 0,04}{0,6} = 0,367 (\text{мкг}).$$

Статистическая обработка результатов анализа фотометрического определения Zn^{2+} -ионов.

1. Рассчитываем стандартное отклонение для результата анализа:

$$s_{x_{\text{контр.}}} = \frac{s_0}{b} \left[\frac{1}{m} + \frac{1}{n} + \frac{n(\bar{y}_{\text{контр.}} - \bar{y})^2}{b^2 [n \sum x_i^2 - (\sum x_i)^2]} \right]^{1/2},$$

$$s_{x_{\text{контр.}}} = \frac{2,8 \cdot 10^{-4}}{0,6} \left[\frac{1}{3} + \frac{1}{6} + \frac{6(0,260 - 0,193)^2}{(0,6)^2 [6 \cdot 0,55 - (1,50)^2]} \right]^{1/2} = \frac{2,8 \cdot 10^{-4} \cdot 0,755}{0,6} = 3,5 \cdot 10^{-4}.$$

2. Рассчитываем доверительный интервал значения результата анализа:

$$\Delta \bar{x}_{\text{контр.}} = t_{P, f} \cdot s_{x_{\text{контр.}}};$$

$$\Delta \bar{x}_{\text{контр.}} = t_{0,95; 4} \cdot 3,5 \cdot 10^{-4};$$

$$\Delta \bar{x}_{\text{контр.}} = 2,78 \cdot 3,597 \cdot 10^{-4} \approx 0,001 \text{ (мкг)}.$$

3. Рассчитываем интервальные значения содержания определяемого компонента:

$$\bar{x}_{\text{контр.}} \pm \Delta \bar{x}_{\text{контр.}} = 0,360 \pm 0,001 = 0,359 \div 0,361 \text{ (мкг)}.$$

4. Рассчитываем относительную погрешность результата анализа (%):

$$\frac{\Delta \bar{x}}{\bar{x}} \cdot 100 = \frac{0,001 \cdot 100}{0,360} = 0,28.$$

5. Полученные результаты заносим в таблицу (таблица 8).

Таблица 8 – Результаты фотометрического определения содержания Zn^{2+} -ионов в контрольной пробе

$A_{\text{контр.}}$	\bar{A}	$\bar{m}(\text{Zn}^{2+})_{\text{контр.}}$, мкг	$s_{x_{\text{контр.}}}$, мкг	$\pm \Delta m(\text{Zn}^{2+})$, мкг	$\bar{m}(\text{Zn}^{2+})_{\text{контр.}}$, $\pm \Delta m(\text{Zn}^{2+})$, мкг
0,255 0,260 0,265	0,260	0,360	$3,5 \cdot 10^{-4}$	0,001	0,359 ÷ 0,361

Ответ: При доверительной вероятности 0,95 и при отсутствии систематических погрешностей действительное значение массы Zn^{2+} -ионов в контрольной пробе находится в интервале от 0,359 до 0,361 мкг.

3.3 Проверка линейности градуировочного графика

Линейная зависимость y от x при построении градуировочного графика нередко не выполняется, поэтому приходится ее проверять. Для этого надо так планировать эксперимент, чтобы при нахождении градуировочной зависимости для каждого стандарта x_i было выполнено m параллельных определений аналитического сигнала y_i и найдено среднее значение \bar{y}_i из m параллельных определений.

Линейность зависимости проверяется путем сравнения дисперсии (s_o^2) линии регрессии, рассеяния средних значений (\bar{y}_i) с дисперсией воспроизводимости $s_{воспр.}^2$, обусловленной рассеянием результатов y_i параллельных определений относительно среднего значения \bar{y}_i в каждой экспериментальной точке.

Дисперсию s_o^2 находят по формуле:

$$s_o^2 = \frac{m}{n-2} \sum_{i=1}^n (\bar{y}_i - Y_i)^2, \quad (51)$$

где \bar{y}_i – среднее значение аналитического сигнала для i -серии из m параллельных определений;

Y_i – вычисленное значение аналитического сигнала для каждого отдельного опыта i -серии (для каждой отдельной точки градуировочного графика;

n – число точек, взятых для нахождения зависимости (число измерений).

Дисперсию воспроизводимости $s_{воспр.}^2$, вычисляют по формуле:

$$s_{воспр.}^2 = \frac{s_1^2 + s_2^2 + \dots + s_n^2}{n} = \frac{1}{n} \sum_{i=1}^n s_i^2 = \frac{1}{n} \cdot \frac{\sum_{j=1}^m (y_{1j} - \bar{y}_1)^2}{m-1} +$$

$$+ \frac{1}{n} \cdot \frac{\sum_{j=1}^m (y_{2j} - \bar{y}_2)^2}{m-1} + \dots + \frac{1}{n} \cdot \frac{\sum_{j=1}^m (y_{nj} - \bar{y}_n)^2}{m-1} = \frac{\sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \bar{y}_i)^2}{n(m-1)}, \quad (52)$$

где $s_i^2 = \frac{\sum_{j=1}^m (y_{ij} - \bar{y}_i)^2}{m-1}$ – дисперсия i -серии;

$y_{1j}; y_{2j}; \dots y_{nj}$; – j -значение аналитического сигнала соответственно в 1, 2, ..., n серии;

$\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n$ – среднее значение аналитического сигнала соответственно в 1, 2, ..., n серии измерений.

Очевидно, что значения дисперсий s_o^2 и $s_{воспр.}^2$ не совпадают. Необходимо выяснить, является ли отличие дисперсий значимым или имеет случайный характер. Эта задача решается с помощью распределения Фишера.

$$\text{Отношение } F_{\text{эсп.}} = \frac{s_{(1)}^2}{s_{(2)}^2}$$

должно быть меньше, чем $F_{\text{табл.}}$ при числе степеней свободы $f = n-2$ для дисперсии s_o^2 и $f = n(m-1)$ для дисперсии $s_{воспр.}^2$. Если $F_{\text{эсп.}} > F_{\text{табл.}}$, то отличие дисперсий s_o^2 и $s_{воспр.}^2$ значимо и гипотеза линейности графика не принимается.

Характер расположения экспериментальных точек градуировочной зависимости часто указывает на то, что график лучше описывается уравнением параболы

$$Y = ax + bx^2, \quad (53)$$

проходящей через начало координат.

Для вычисления оптимальных значений параметров a и b квадратичной градуировочной зависимости (53) метод наименьших квадратов дает следующие формулы:

$$a = \frac{\sum x_i y_i \cdot \sum x_i^4 - \sum x_i^3 \cdot \sum x_i^2 y_i}{\sum x_i^4 \cdot \sum x_i^2 - (\sum x_i^3)^2}; \quad (54)$$

$$b = \frac{\sum x_i^2 y_i \cdot \sum x_i^2 - \sum x_i y_i \cdot \sum x_i^3}{\sum x_i^4 \cdot \sum x_i^2 - (\sum x_i^3)^2} . \quad (55)$$

3.4 Контрольные вопросы

1. Пояснить смысл параметров a и b в уравнении прямой $Y = a + bx$.
2. Перечислить условия графического определения a и b .
3. Назвать способы построения «наилучшей» возможной прямой.
4. Описать графический способ построения «наилучшей» прямой по результатам измерений.
5. Дать понятие «эмпирические формулы».
6. Пояснить сущность метода наименьших квадратов.
7. Пояснить смысл записи $\sum_{i=1}^n \delta_{y_i}^2$.
8. Дать определение нормальной системы уравнений.
9. Перечислить известные методы решения системы n линейных уравнений с n неизвестными.
10. Пояснить применение метода Гаусса для решения системы n линейных уравнений с n неизвестными ($n \times n$).
11. Привести формулы Крамера для системы $(n \times n)$ линейных уравнений.
12. Сформулировать условие экстремума функции $SQ(a, b, c, \dots)$.
13. Сформулировать задачу метода наименьших квадратов для линейной зависимости в аналитическом виде.
14. Вывести для функции $SQ = \sum (a + bx_i - y_i)^2$ нормальную систему уравнений.
15. Записать формулы Крамера для нахождения неизвестных параметров a и b из нормальной системы уравнений.

16. Записать формулы для определения неизвестных параметров a и b уравнения прямой $Y = a + bx$, выведенных по методу наименьших квадратов.

17. Обосновать необходимость статистической обработки параметров градуировочного графика.

18. Написать формулу для определения дисперсии линии регрессии s_0^2 .

19. Показать сущность метода Гаусса для определения коэффициентов при квадратичной зависимости.

20. Описать проверку линейности градуировочного графика.

4 Применение метода наименьших квадратов для обработки нелинейных зависимостей

В физико-химических исследованиях часто наблюдаются нелинейные функциональные зависимости, для обработки которых также возможно применение метода наименьших квадратов.

Рассмотрим возможности применения метода наименьших квадратов для наиболее распространенных эмпирических зависимостей: квадратичной, показательной и степенной.

4.1 Обработка функций квадратичной зависимости

Аналитически задача метода наименьших квадратов для квадратичной функции вида

$$Y = a + bx + cx^2 \quad (56)$$

формулируется следующим образом:

$$\sum \delta_{y_i}^2 = \sum (Y_i - y_i)^2 = \sum [(a + bx_i + cx_i^2) - y_i]^2 \rightarrow \min . \quad (57)$$

Учитывая обозначение $\sum \delta_{y_i}^2$ как SQ , получим:

$$SQ = \sum [(a + bx_i + cx_i^2) - y_i]^2 . \quad (58)$$

Функция SQ – функция трех переменных a , b и c , а x_i и y_i – постоянные числа, найденные экспериментально. Аналогично изложенному выше (раздел 3.2), минимум этой функции определяется из системы нормальных уравнений:

$$\begin{cases} SQ'_a = \frac{\partial SQ}{\partial a} = 0 \\ SQ'_b = \frac{\partial SQ}{\partial b} = 0 \\ SQ'_c = \frac{\partial SQ}{\partial c} = 0 \end{cases} . \quad (59)$$

Найдем частные производные первого порядка для функции (58)

$$SQ = \sum [(a + bx_i + cx_i^2) - y_i]^2$$

по переменным a , b , c и запишем систему уравнений:

$$\begin{cases} SQ'_a = \frac{\partial SQ}{\partial a} = 2 \sum (a + bx_i + cx_i^2 - y_i) = 2(an + b \sum x_i + c \sum x_i^2 - \sum y_i) \\ SQ'_b = \frac{\partial SQ}{\partial b} = 2 \sum (a + bx_i + cx_i^2 - y_i) \cdot x_i = 2(a \sum x_i + b \sum x_i^2 + c \sum x_i^3 - \sum x_i y_i) \\ SQ'_c = \frac{\partial SQ}{\partial c} = 2 \sum (a + bx_i + cx_i^2 - y_i) \cdot x_i^2 = 2(a \sum x_i^2 + b \sum x_i^3 + c \sum x_i^4 - \sum x_i^2 y_i) . \end{cases} \quad (60)$$

После сокращения всех уравнений на 2 эта система нормальных уравнений примет вид:

$$\begin{cases} an + b \sum x_i + c \sum x_i^2 - \sum y_i = 0 \\ a \sum x_i + b \sum x_i^2 + c \sum x_i^3 - \sum x_i y_i = 0 \\ a \sum x_i^2 + b \sum x_i^3 + c \sum x_i^4 - \sum x_i^2 y_i = 0 \end{cases} \quad (61)$$

$$\text{или} \quad \begin{cases} an + b \sum x_i + c \sum x_i^2 = \sum y_i \\ a \sum x_i + b \sum x_i^2 + c \sum x_i^3 = \sum x_i y_i \\ a \sum x_i^2 + b \sum x_i^3 + c \sum x_i^4 = \sum x_i^2 y_i \end{cases} . \quad (62)$$

Систему (62) можно решить по правилу Крамера или методом Гаусса в общем виде и найти a , b и c , но при решении получим довольно громоздкие выкладки. Поэтому лучше и проще для конкретной задачи записать нормальную систему трех уравнений с тремя неизвестными a , b , c , а указанные в системе (62)

суммы будут определенными числами и тогда уже полученную систему решить одним из предложенных методов, то есть найти a , b и c .

Пример 4.1 На основании значений теплоемкости метана при нескольких температурах выразить эмпирическим уравнением зависимость теплоемкости метана от температуры.

T, K	300	400	500	600	700	800	900	1000
$C_p^o, \text{Дж}/(\text{моль}\cdot\text{К})$	35,80	40,74	46,56	52,50	58,07	63,27	67,91	72,06

Решение

Зависимость теплоемкости от температуры для органических веществ выражается эмпирическим уравнением параболы

$$C_p^o = A + BT + CT^2, \quad (63)$$

где C_p^o — теплоемкость вещества при постоянном давлении в стандартных условиях.

Для определения коэффициентов A , B , и C применим метод наименьших квадратов.

Обозначим:

$$C_p^o = Y; \quad \frac{T-300}{100} = x,$$

тогда получим квадратичную функцию вида:

$$Y = a + bx + cx^2. \quad (64)$$

Коэффициенты a , b и c не равны коэффициентам A , B , C , так как за x обозначено не T , а $\frac{T-300}{100}$ (для упрощения расчетов).

В качестве наилучших значений параметров a , b , и c следует выбрать такие, для которых сумма квадратов отклонений экспериментальных значений от вычисленных ($\delta_{y_i}^2$) минимальна.

Для нахождения сумм, указанных в системе (62), составим таблицу 9.

Таблица 9 - Исходные данные и вспомогательные расчеты при определении оптимальных параметров a , b и c методом наименьших квадратов

T, K	$C_{P_i \text{ эксп.}}^o = y_i$	x_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y_i$	$C_{P_i \text{ расч.}}^o = Y_i$
300	35,80	0	0	0	0	0	0	35,28
400	40,74	1	1	1	1	40,74	40,74	41,25
500	46,56	2	4	8	16	93,12	186,24	46,99
600	52,50	3	9	27	81	157,50	472,50	52,52
700	58,07	4	16	64	256	232,28	929,12	57,82
800	63,27	5	25	125	625	316,35	1581,75	62,90
900	67,91	6	36	216	1296	407,46	2444,76	67,76
1000	72,06	7	49	343	2401	504,42	3530,94	72,39

$$\sum_{i=1}^8 y_i = 436,91; \quad \sum_{i=1}^8 x_i = 28; \quad \sum_{i=1}^8 x_i^2 = 140; \quad \sum_{i=1}^8 x_i^3 = 784;$$

$$\sum_{i=1}^8 x_i^4 = 4676; \quad \sum_{i=1}^8 x_i y_i = 1751,87; \quad \sum_{i=1}^8 x_i^2 y_i = 9186,05.$$

Подставляя в уравнение системы (62) значения сумм, получим:

$$\begin{cases} 8a + 28b + 140c = 436,91 & (a') \\ 28a + 140b + 784c = 1751,87 & (b') \\ 140a + 784b + 4676c = 9186,05 & (c') \end{cases} \quad (65)$$

Решим систему уравнений (65) методом Гаусса. Для этого умножим уравнение $(65 a')$ на 7, а $(65 b')$ на 2 и вычтем из полученного второго уравнения первое:

$$\begin{cases} 56a + 196b + 980c = 3058,37 \\ \underline{56a + 280b + 1568c = 3503,74} \end{cases} \quad (66)$$

$$84b + 588c = 445,37.$$

Аналогично, умножим уравнение $(65 b')$ на 5 и полученное уравнение вычтем из уравнения $65 c'$:

$$\begin{cases} 140a + 700b + 3920c = 8759,35 \\ \underline{140a + 784b + 4676c = 9186,05} \end{cases} \quad (67)$$

$$84b + 756c = 426,80.$$

Из полученной системы уравнений с двумя неизвестными найдем c :

$$\begin{cases} 84b + 588c = 445,37 \\ \underline{84b + 756c = 426,80} \end{cases} \quad (68)$$

$$168c = -18,67,$$

$$c = -0,1111.$$

Из первого уравнения системы (68) находим b :

$$b = \frac{445,37 - 588(-0,1111)}{84} = 6,0797.$$

Подставляя полученные значения коэффициентов b и c в первое уравнение системы (65), находим a :

$$a = \frac{3058,37 - 196 \cdot 6,0797 - 980 \cdot (-0,1111)}{56} = \frac{3058,37 - 1191,62 + 108,878}{56} = 35,2791.$$

Начальное уравнение зависимости теплоемкости метана от температуры имеет вид:

$$C_p^o = 35,2791 + 6,0797 \left(\frac{T-300}{100} \right) - 0,1111 \left(\frac{T-300}{100} \right)^2. \quad (69)$$

Преобразуем полученное уравнение и найдем коэффициенты A , B и C , а также запишем эмпирическое уравнение зависимости теплоемкости метана от температуры.

$$C_p^o = 35,2791 + 6,0797 \cdot 10^{-2} T - 18,2391 - 0,1111 \cdot 10^{-4} T^2 + 0,6666 \cdot 10^{-2} T - 0,9999 = \\ = 16,040 + 67,463 \cdot 10^{-3} T - 11,11 \cdot 10^{-6} T^2.$$

$$A = 16,040; \quad B = 67,463 \cdot 10^{-3}; \quad C = -11,11 \cdot 10^{-6}.$$

$$C_p^o = 16,040 + 67,463 \cdot 10^{-3} T - 11,11 \cdot 10^{-6} T^2. \quad (70)$$

Значения C_p^o , вычисленные по этой формуле ($C_{P_i}^o \text{ расч.}$) для восьми значений температуры, даны в последнем столбце таблицы 9.

Уравнение (70) можно использовать для вычисления любых промежуточных значений теплоемкости метана. Графически зависимость теплоемкости метана от температуры в исследуемом интервале выражается частью параболы, показанной на рисунке 8, построенной по рассчитанным значениям теплоемкости ($C_{P_i}^o \text{ расч.}$).

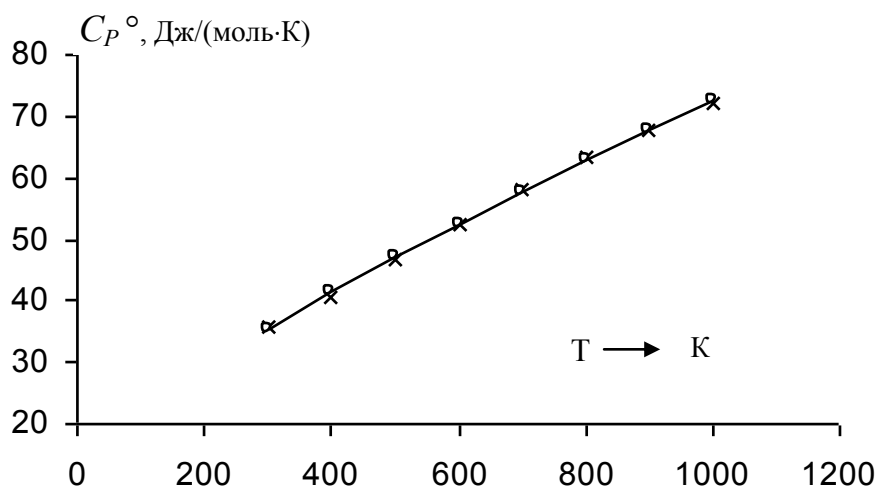


Рисунок 8 - Зависимость теплоемкости метана от температуры;
 \times — экспериментальные значения; \circ — рассчитанные значения

4.2 Обработка функций показательной зависимости

Показательная зависимость имеет вид:

$$y = a \cdot e^{bx} . \quad (71)$$

Нахождение параметров a и b показательной зависимости методом наименьших квадратов производится следующим образом. Прологарифмируем по основанию 10 обе части формулы (70). Получим:

$$\lg y = \lg a + bx \cdot \lg e . \quad (72)$$

Обозначим $\lg e = M$ ($\lg e = 0,4343$); $\lg y = Y$, тогда

$$Y = \lg a + bxM . \quad (73)$$

Для данной функции сумма квадратов отклонений будет иметь вид:

$$\sum \delta_{y_i}^2 = \sum (Y_i - y_i)^2 = \sum (\lg a + bx_i M - y_i)^2 ,$$

где $\delta_{y_i} = \delta(a, b)$ – функция двух переменных;

x_i и y_i – постоянные числа, найденные экспериментально.

Учитывая обозначение $\sum \delta_{y_i}^2$ как SQ , получим:

$$SQ = \sum (\lg a + bx_i M - y_i)^2 .$$

Минимум этой функции определяется по правилам дифференциального исчисления функции двух переменных из системы

$$\begin{cases} SQ'_a = \frac{\partial SQ}{\partial a} = 0 \\ SQ'_b = \frac{\partial SQ}{\partial b} = 0 \end{cases},$$

где SQ'_a и SQ'_b – частные производные первого порядка по переменным a и b , соответственно.

Найдем SQ'_a и SQ'_b и запишем нормальную систему, из которой найдем параметры a и b для уравнения (72).

$$\begin{aligned} SQ'_a &= 2 \sum (\lg a + bx_i M - y_i) \cdot \frac{1}{a} \cdot \lg e = 2 \sum (\lg a + bx_i M - y_i) \cdot \frac{1}{a} \cdot M = \\ &= 2 \left(\sum \lg a \cdot \frac{1}{a} \cdot M + bM^2 \cdot \frac{1}{a} \sum x_i - \sum \frac{1}{a} M y_i \right) = \\ &= 2 \left(\lg a \cdot \frac{1}{a} \cdot M \cdot n + \frac{b}{a} M^2 \sum x_i - \frac{1}{a} \cdot M \sum y_i \right). \\ SQ'_b &= 2 \sum (\lg a + bx_i M - y_i) \cdot x_i M = 2 \sum (\lg a x_i M + \sum bx_i^2 M^2 - \sum y_i x_i M) = \\ &= 2 (M \lg a \sum x_i + bM^2 \sum x_i^2 - M \sum x_i y_i). \end{aligned}$$

Нормальная система уравнений будет иметь вид:

$$\begin{cases} \lg a \cdot \frac{1}{a} \cdot M \cdot n + \frac{b}{a} M^2 \sum x_i = \frac{1}{a} M \sum y_i \\ M \lg a \sum x_i + bM^2 \sum x_i^2 = M \sum x_i y_i \end{cases}, \quad (74)$$

так как $M \neq 0$ и $\frac{1}{a} \neq 0$, то, поделив первое уравнение системы (74) на $\frac{1}{a} \cdot M$, а

второе – на M , получим систему:

$$\begin{cases} \lg a \cdot n + bM \sum x_i = \sum y_i \\ \lg a \sum x_i + bM \sum x_i^2 = \sum x_i y_i \end{cases} \quad (75)$$

или

$$\begin{cases} bM \sum x_i + \lg a \cdot n = \sum y_i \\ bM \sum x_i^2 + \lg a \sum x_i = \sum x_i y_i \end{cases} . \quad (76)$$

Помножим первое уравнение на $\sum x_i$, а второе – на n и вычтем из второго уравнения первое.

$$\begin{aligned} & bM (\sum x_i)^2 + n \lg a \sum x_i = \sum x_i \sum y_i \\ & \underline{bn \cdot M \sum x_i^2 + n \lg a \sum x_i = n \sum x_i y_i} \\ & bnM \sum x_i^2 - bM (\sum x_i)^2 = n \sum x_i y_i - \sum x_i \sum y_i \\ & b[nM \sum x_i^2 - M (\sum x_i)^2] = n \sum x_i y_i - \sum x_i \sum y_i \end{aligned}$$

Отсюда находим b :

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{M[n \sum x_i^2 - (\sum x_i)^2]} . \quad (77)$$

Подставим полученное значение b в первое уравнение системы (76) и определим $\lg a$.

$$\begin{aligned} n \lg a &= \sum y_i - bM \sum x_i ; \\ n \lg a &= \sum y_i - \frac{(n \sum x_i y_i - \sum x_i \sum y_i)}{M[n \sum x_i^2 - (\sum x_i)^2]} \cdot M \sum x_i = \\ &= \frac{n \sum y_i \sum x_i^2 - \sum y_i (\sum x_i)^2 - n \sum x_i \sum x_i y_i + (\sum x_i)^2 \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \\ &= \frac{n \sum y_i \sum x_i^2 - n \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{n(\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i)}{n \sum x_i^2 - (\sum x_i)^2} . \\ n \lg a &= \frac{n(\sum y_i \sum x_i^2 - \sum x_i y_i \sum x_i)}{n \sum x_i^2 - (\sum x_i)^2} , \text{ тогда} \\ \lg a &= \frac{\sum y_i \sum x_i^2 - \sum x_i y_i \sum x_i}{n \sum x_i^2 - (\sum x_i)^2} , \text{ отсюда} \end{aligned}$$

$$a = 10^{\frac{\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i}{n \sum x_i^2 - (\sum x_i)^2}}. \quad (78)$$

Таким образом, для показательной функции $y = a \cdot e^{bx}$ параметры a и b вычисляются по формулам:

$$a = 10^{\frac{\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i}{n \sum x_i^2 - (\sum x_i)^2}}; \quad b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{0,4343 [n \sum x_i^2 - (\sum x_i)^2]}.$$

Аналогично нахождению параметров a и b для линейной функции, при данных значениях a и b функция

$$SQ = \sum (\lg a + bx_i M - y_i)^2$$

принимает минимальное значение.

Напомним, что при выводе формул (77) и (78) за Y обозначен $\lg b$ (логарифм по основанию 10).

Пример 4.2 - Зависимость константы скорости реакции от абсолютной температуры выражается показательной функцией вида:

$$k = A e^{-\frac{E}{RT}}.$$

Опытным путем определены константы скорости реакции (k) при шести различных температурах (T):

T, K	673	725	766	801	834	877
k	3,23	7,80	15,43	24,21	37,95	60,09

Найти наиболее вероятные численные значения множителя A , стоящего перед экспонентой, и E/R .

Решение

Способ 1

Для расчета коэффициентов A и $\frac{E}{R}$ в исходном уравнении Аррениуса

$$k = A e^{-\frac{E}{RT}} \quad (79)$$

прологарифмируем (по основанию 10) обе части его. Получим:

$$\lg k = \lg A - \frac{E}{R} \cdot \frac{1}{T} \lg e, \quad \text{где } \lg e \approx 0,4343. \quad (80)$$

Обозначим $\lg k = Y, \quad A = a, \quad -\frac{E}{R} = b, \quad \frac{1}{T} = x.$

При этом линейную функцию $\lg k = f(1/T)$ приводим к стандартной линейной зависимости

$$Y = \lg a + bx \lg e,$$

коэффициенты a и b которой определяются по формулам (79, 80).

Исходные данные и вспомогательные расчеты для определения коэффициентов a и b сведем в таблицу 10.

Таблица 10 – Исходные данные и вспомогательные расчеты при определении значений параметров a и b по методу наименьших квадратов

k_i	T_i	$\lg k_i = y_i$	$1/T_i = x_i$	x_i^2	$x_i y_i$	$k_{i \text{ расч.}}$
3,23	673	0,5092	$1,486 \cdot 10^{-3}$	$2,208 \cdot 10^{-6}$	$0,752 \cdot 10^{-3}$	3,35
7,80	725	0,8921	$1,379 \cdot 10^{-3}$	$1,902 \cdot 10^{-6}$	$1,230 \cdot 10^{-3}$	8,13
15,43	766	1,1884	$1,306 \cdot 10^{-3}$	$1,706 \cdot 10^{-6}$	$1,552 \cdot 10^{-3}$	15,03
24,21	801	1,3840	$1,248 \cdot 10^{-3}$	$1,558 \cdot 10^{-6}$	$1,727 \cdot 10^{-3}$	24,16
37,95	834	1,5792	$1,199 \cdot 10^{-3}$	$1,438 \cdot 10^{-6}$	$1,894 \cdot 10^{-3}$	36,44
60,09	877	1,7788	$1,140 \cdot 10^{-3}$	$1,300 \cdot 10^{-6}$	$2,028 \cdot 10^{-3}$	59,43

$$\sum y_i = 7,3318; \quad \sum x_i = 7,758 \cdot 10^{-3}; \quad \sum x_i^2 = 10,112 \cdot 10^{-6}; \quad \sum x_i y_i = 9,1876 \cdot 10^{-3}.$$

$$\begin{aligned} a &= 10^{\frac{\sum x_i^2 \sum y_j - \sum x_i y_j \sum x_j}{n \sum x_j^2 - (\sum x_j)^2}} = 10^{\frac{10,112 \cdot 10^{-6} \cdot 7,3318 - 9,1876 \cdot 10^{-3} \cdot 7,758 \cdot 10^{-3}}{6 \cdot 10,112 \cdot 10^{-6} - (7,758 \cdot 10^{-3})^2}} = \\ &= 10^{\frac{2,8618 \cdot 10^{-6}}{0,4854 \cdot 10^{-6}}} = 10^{5,89567} = 7,864 \cdot 10^5. \end{aligned}$$

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{0,4343 [n \sum x_i^2 - (\sum x_i)^2]} = \frac{6 \cdot 9,1876 \cdot 10^{-3} - 7,758 \cdot 10^{-3} \cdot 7,3318}{0,4343 [6 \cdot 10,112 \cdot 10^{-6} - (7,758 \cdot 10^{-3})^2]} =$$

$$= -\frac{1,7545 \cdot 10^{-3}}{0,4343 \cdot 0,4854 \cdot 10^{-6}} = -8322,69 \approx -8323.$$

Находим множитель A , стоящий перед экспонентой, и $\frac{E}{R}$.

$$A = a = 7,864 \cdot 10^5 ;$$

$$-\frac{E}{R} = b = -8323; \quad \frac{E}{R} = 8323.$$

Таким образом, зависимость константы скорости исследуемой реакции от температуры выражается эмпирическим уравнением

$$k = 7,864 \cdot 10^5 \cdot e^{-\frac{8323}{T}}.$$

Значения k , вычисленные по этой формуле ($k_{расч.}$) для шести значений температуры, даны в последнем столбце таблицы 10.

Зависимость константы скорости реакции от температуры в исследуемом интервале температур выражается частью графика экспоненциальной функции (рисунок 9), построенного по рассчитанным значениям константы скорости ($k_{i расч.}$).

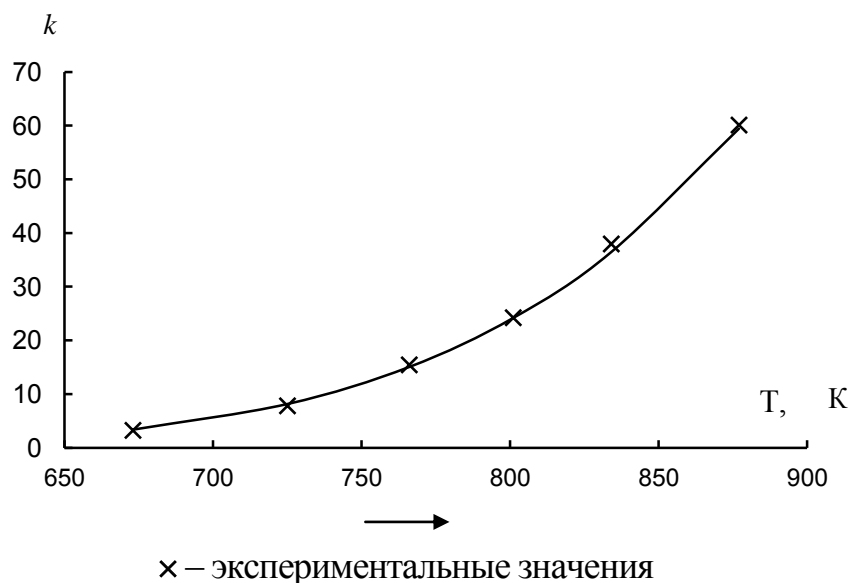


Рисунок 9 - Зависимость константы скорости реакции от температуры

Способ 2

Не всегда целесообразно использовать формулы (77) и (78) для нахождения параметров a и b в показательной функции $y = ae^{bx}$. Как известно, при логарифмировании данной зависимости по основанию 10 функцию (72) приводим к линейному виду :

$$\lg b = \lg a + bx \cdot \lg e.$$

Тогда нормальная система будет иметь вид:

$$\begin{cases} b \lg e \sum x_i + \lg a \cdot n = \sum \lg y_i \\ b \lg e \sum x_i^2 + \lg a \sum x_i = \sum x_i \lg y_i \end{cases} \quad (81)$$

Это система двух линейных уравнений с двумя неизвестными a и b , которую необходимо решить по методу Гаусса или по формулам Крамера. Рассмотрим этот способ решения на примере уравнения Аррениуса:

$$k = A e^{-\frac{E}{RT}},$$

используя следующие обозначения:

$$k = y, \quad A = a, \quad -\frac{E}{R} = b, \quad \frac{1}{T} = x.$$

Исходные данные и вспомогательные расчеты для определения коэффициентов a и b в системе возьмем из таблицы 10. Тогда система уравнений (81) будет иметь следующий вид:

$$\begin{cases} b \lg e \cdot 7,758 \cdot 10^{-3} + 6 \lg a = 7,3318 \\ b \lg e \cdot 10,112 \cdot 10^{-6} + \lg a \cdot 7,758 \cdot 10^{-3} = 9,1876 \cdot 10^{-3} \end{cases} \quad (82)$$

Упростим систему, умножив первое уравнение – на 10^3 , а второе – на 10^6 . Получим систему:

$$\begin{cases} b \lg e \cdot 7,758 + 6 \cdot 10^3 \lg a = 7,3318 \cdot 10^3 \\ b \lg e \cdot 10,112 + \lg a \cdot 7,758 \cdot 10^3 = 9,1876 \cdot 10^3 \end{cases}$$

Поскольку $\lg e = 0,4343$, получим:

$$\begin{cases} 0,4343 \cdot 7,758b + 6000 \lg a = 7331,8 \\ 0,4343 \cdot 10,112b + 7758 \lg a = 9187,6 \end{cases} \quad (83)$$

Умножим первое уравнение системы (83) на $-10,112$, а второе – на $7,758$ и почленно сложим их:

$$\begin{cases} -0,4343 \cdot 7,758 \cdot 10,112b - 6000 \cdot 10,112 \lg a = -7331,8 \cdot 10,112 \\ 0,4343 \cdot 10,112 \cdot 7,758b + 7758 \cdot 7,758 \lg a = 9187,6 \cdot 7,758 \end{cases}$$

Получим уравнение с одним неизвестным ($\lg a$):

$$7758 \cdot 7,758 \lg a - 10,112 \cdot 6000 \lg a = 9187,6 \cdot 7,758 - 7331,8 \cdot 10,112;$$

$$-485,436 \lg a = 2861,7608;$$

$$\lg a = 5,8953.$$

$$a = 10^{5,8953} = 7,858 \cdot 10^5.$$

Из первого уравнения системы (83) получим:

$$0,4343 \cdot 7,758b = 7331,8 - 6000 \lg a;$$

$$\text{отсюда} \quad b = \frac{7331,8 - 5,8953 \cdot 6000}{0,4343 \cdot 7758} = -8322.$$

$$\text{Находим } A \text{ и } \frac{E}{R}. \quad A = 7,858 \cdot 10^5; \quad \frac{E}{R} = 8322.$$

Способ 3

Учитывая, что при логарифмировании по основанию 10 функции $y = a \cdot e^{bx}$,

показательная зависимость превращается в линейную $\lg y = \lg a + bx \cdot \lg e$, оптимальные параметры $\lg a$ и $b \cdot \lg e$ можно найти с помощью стандартных формул, выведенных для линейной функции

$$Y = a + bx.$$

Обозначим $\lg y = Y$; $\lg a = a'$; $b \cdot \lg e = b'$. Тогда уравнение будет иметь вид:

$$Y = a' + b'x, \quad (84)$$

коэффициенты a' и b' которого находим по формулам:

$$a' = \frac{\sum x_i^2 \cdot \sum y_i - \sum x_i \cdot \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2};$$

$$b' = \frac{n \sum x_i y_i - \sum x_i \cdot \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}.$$

Для решения данным способом задачи, рассмотренной выше, исходные данные и вспомогательные расчеты возьмем из таблицы 10. Получим:

$$a' = \frac{10,112 \cdot 10^{-6} \cdot 7,3318 - 9,1876 \cdot 10^{-3} \cdot 7,758 \cdot 10^{-3}}{6 \cdot 10,112 \cdot 10^{-6} - (7,758 \cdot 10^{-3})^2} = 5,89567,$$

$$b' = \frac{6 \cdot 9,1876 \cdot 10^{-3} - 7,758 \cdot 10^{-3} \cdot 7,3318}{6 \cdot 10,112 \cdot 10^{-6} - (7,758 \cdot 10^{-3})^2} = -3,6145.$$

Поскольку $\lg a = a'$, отсюда

$$a = 10^{5,89567} = 7,864 \cdot 10^5,$$

$$b \cdot \lg e = b', \lg e = 0,4343.$$

$$\text{Тогда} \quad b = -\frac{3,6145}{0,4343} = -8323.$$

Находим A и $\frac{E}{R}$.

$$A = a = 7,864 \cdot 10^5; \quad -\frac{E}{R} = b = -8323; \quad \frac{E}{R} = 8323.$$

Способ 4

Выравнивание показательной зависимости можно провести также путем логарифмирования по основанию e обеих частей уравнения (84). Получим:

$$\ln k = \ln A - \frac{E}{R} \cdot \frac{1}{T}. \quad (85)$$

Обозначим $\ln k = Y$; $\ln A = a$; $-\frac{E}{R} = b$; $\frac{1}{T} = x$, тогда линейную функцию приведем к стандартной линейной зависимости

$$Y = a + bx,$$

Все необходимые вычисления представлены в таблице 11.

Таблица 11 - Исходные данные и вспомогательные расчеты при определении оптимальных параметров a и b по методу наименьших квадратов

k_i	T_i	$\ln k_i = y_i$	$\frac{1}{T_i} \cdot 10^3 = x_i$	$x_i^2 \cdot 10^6$	$x_i y_i \cdot 10^3$	$k_{i \text{ расч.}}$	$\ln k_{i \text{ расч.}} = Y$
3,23	673	1,1725	1,486	2,208	1,742	3,35	1,2090
7,80	725	2,0541	1,379	1,902	2,833	8,13	2,0956
15,43	766	2,7363	1,306	1,706	3,574	15,03	2,7100
24,21	801	3,1868	1,248	1,558	3,977	24,16	3,1847
37,95	834	3,6363	1,199	1,438	4,360	36,44	3,5957
60,09	877	4,0958	1,140	1,300	4,669	59,43	4,0848

$$\sum y_i = 16,8818; \quad \sum x_i = 7,758 \cdot 10^{-3}; \quad \sum x_i^2 = 10,112 \cdot 10^{-6}; \quad \sum x_i y_i = 21,155 \cdot 10^{-3}.$$

$$a = \frac{\sum x_i^2 \cdot \sum y_i - \sum x_i \cdot \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{10,112 \cdot 10^{-6} \cdot 16,8818 - 7,758 \cdot 10^{-3} \cdot 21,155 \cdot 10^{-3}}{6 \cdot 10,112 \cdot 10^{-6} - (7,758 \cdot 10^{-3})^2} =$$

$$= \frac{6,5889 \cdot 10^{-6}}{0,4854 \cdot 10^{-6}} = 13,5728.$$

$$b = \frac{n \sum x_i y_i - \sum x_i \cdot \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{6 \cdot 21,155 \cdot 10^{-3} - 7,758 \cdot 10^{-3} \cdot 16,8818}{0,4854 \cdot 10^{-6}} =$$

$$= -\frac{4,039 \cdot 10^{-3}}{0,4854 \cdot 10^{-6}} = -8320,97 \approx -8321.$$

Находим A и $\frac{E}{R}$.

$$\ln A = a = 13,5728; \quad A = e^{13,5728} = 7,845 \cdot 10^5.$$

$$-\frac{E}{R} = b = -8321 ; \quad \frac{E}{R} = 8321 .$$

Эмпирическая формула зависимости константы скорости исследуемой реакции от температуры имеет вид:

$$k = 7,845 \cdot 10^5 \cdot e^{-\frac{8321}{T}} .$$

Значения k , вычисленные по этой формуле ($k_{\text{расч.}}$) для шести значений температуры, и значения $\ln k_{\text{расч.}}$ даны в последних столбцах таблицы 11.

Зависимость логарифма константы скорости реакции ($\ln k$) от температуры $1/T$ представлена на рисунке 10 графиком, построенным по вычисленным значениям $\ln k_{\text{расч.}}$. Эта зависимость выражается прямой линией.

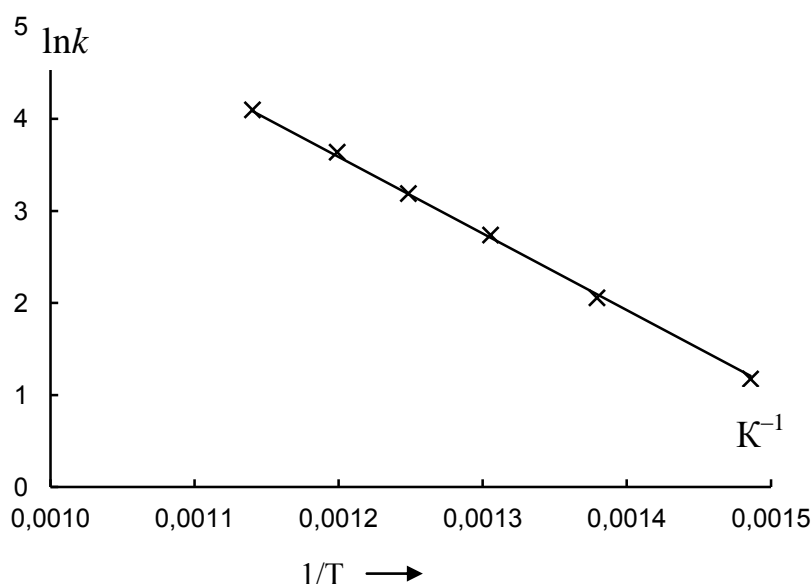


Рисунок 10 - График зависимости логарифма константы скорости реакции от обратной температуры, × – экспериментальные значения

Незначительные расхождения значений констант A и $\frac{E}{R}$ при расчетах по способам 1- 4 связаны с округлением результатов в промежуточных расчетах.

При вычислениях все расчеты необходимо выполнять с бóльшим числом знаков после запятой. Округления в промежуточных расчетах могут привести к бóльшим погрешностям при вычислениях.

Наряду с показательной функцией *с основанием e* вида $y = a \cdot e^{bx}$ встречается и показательная функция *с основанием b* вида $y = a \cdot b^x$.

В том случае, когда имеется показательная функция с основанием b вида $y = a \cdot b^x$, для выравнивания показательной зависимости используют также логарифмическое преобразование функции $y = a \cdot b^x$ к виду

$$\lg y = \lg a + x \lg b. \quad (86)$$

Функция (86) $\lg y = f(x)$ является линейной функцией с параметрами $\lg a$ и $\lg b$. Для определения этих параметров методом наименьших квадратов необходимо найти минимум суммы

$$SQ = \sum_{i=1}^n (x_i \lg b + \lg a - \lg y_i)^2 \quad \text{при условиях}$$

$$\frac{\partial SQ}{\partial \lg a} = 0 \quad \text{и} \quad \frac{\partial SQ}{\partial \lg b} = 0.$$

Эти условия приводят к нормальной системе

$$\begin{cases} \lg b \sum x_i^2 + \lg a \sum x_i = \sum x_i \cdot \lg y_i \\ \lg b \sum x_i + n \cdot \lg a = \sum \lg y_i, \end{cases} \quad (87)$$

которая решается методом Крамера или методом Гаусса.

Из найденных при решении этой системы значений $\lg a$ и $\lg b$ находим a и b .

Пример 4.3 - На основании статистических данных составить эмпирическую формулу роста выпуска предприятием минеральных удобрений за 16 лет с 1980 по 1995 гг. в виде показательной функции $y = a \cdot b^x$, где x – порядковый номер года, y – выпуск продукции (%). Выпуск продукции в 1980 году принять за 100 %.

№ п/п года	Год	Выпуск продукции, %		№ п/п года	Год	Выпуск продукции, %
1	1980	100		9	1988	229
2	1981	112		10	1989	247
3	1982	126		11	1990	267
4	1983	136		12	1991	284
5	1984	153		13	1992	300
6	1985	171		14	1993	312
7	1986	190		15	1994	341
8	1987	204		16	1995	364

Способ 1

Как известно, при логарифмировании зависимости $y = a \cdot b^x$ по основанию 10 получаем формулу

$$\lg y = \lg a + x \cdot \lg b . \quad (88)$$

Нормальная система будет иметь вид:

$$\begin{cases} \lg b \sum x_i^2 + \lg a \sum x_i = \sum x_i \cdot \lg y_i \\ \lg b \sum x_i + n \cdot \lg a = \sum \lg y_i \end{cases} . \quad (89)$$

Эту систему двух уравнений с двумя неизвестными a и b решим методом Гаусса. Для нахождения сумм, указанных в системе (89), составим таблицу 12.

Таблица 12 - Исходные данные и вспомогательные расчеты при определении оптимальных значений параметров a и b по методу наименьших квадратов

x_i	x_i^2	y_i эксп.	$\lg y_i$	$x_i \cdot \lg y_i$
1	1	100	2,0000	2,0000
2	4	112	2,0492	4,0984
3	9	126	2,1004	6,3011
4	16	136	2,1395	8,5342
5	25	153	2,1847	10,9235
6	36	171	2,2330	13,3980
7	49	190	2,2788	15,9513
8	64	204	2,3096	18,4770
9	81	229	2,3598	21,2385
10	100	247	2,3927	23,9270
11	121	267	2,4265	26,6915
12	144	284	2,4533	29,4398
13	169	300	2,4771	32,2026
14	196	312	2,4942	34,9182
15	225	341	2,5328	37,9913
16	256	364	2,5611	40,9776

$$\sum x_i = 136; \quad \sum x_i^2 = 1496; \quad \sum \lg y_i = 36,9867; \quad \sum x_i \lg y_i = 327,0699 .$$

Полученные суммы подставим в систему (89). Тогда нормальная система (89) примет вид:

$$\begin{cases} 1496 \cdot \lg b + 136 \lg a = 327,0699 \\ 136 \lg b + 16 \lg a = 36,9867 \end{cases} . \quad (90)$$

Поделим первое уравнение системы (90) на 4, а второе – на 8. Получим систему:

$$\begin{cases} 374 \lg b + 34 \lg a = 81,767475 \\ 17 \lg b + 2 \lg a = 4,6233375 \end{cases} . \quad (91)$$

Умножим второе уравнение системы (91) на -17 и прибавим к первому.

$$+ \begin{cases} 374 \lg b + 34 \lg a = 81,767475 \\ -289 \lg b - 34 \lg a = -78,5967375 \end{cases}$$

$$85 \lg b = 3,1707375$$

$$\lg b = 0,0373028 .$$

$$\text{Отсюда } b \approx 1,090 .$$

Подставим полученное значение b во второе уравнение системы (91) и найдем $\lg a$.

$$\lg a = \frac{4,6233375 - 17 \cdot 0,0373028}{2} = 1,994595 ,$$

$$\text{отсюда } a \approx 98,763 .$$

Таким образом, эмпирическое уравнение роста выпуска предприятием минеральных удобрений имеет вид:

$$b = 98,763 \cdot 1,090^x .$$

Способ 2

При логарифмировании по основанию 10 показательная зависимость $y = a \cdot b^x$ превращается в линейную:

$$\lg y = \lg a + x \lg b . \quad (92)$$

Обозначим $\lg y = Y$; $\lg a = a'$; $\lg b = b'$. Тогда линейная функция (92) примет вид:

$$Y = a' + b'x ,$$

что позволяет найти оптимальные значения параметров a' и b' по выведенным стандартным формулам, учитывая, что $\lg y = Y$. Необходимые суммы для определения коэффициентов a' и b' возьмем из таблицы 13.

$$a' = \frac{\sum x_i^2 \cdot \sum \lg y_i - \sum x_i \cdot \sum x_i \lg y_i}{n \sum x_i^2 - (\sum x_i)^2} ;$$

$$a' = \frac{1496 \cdot 36,9867 - 136 \cdot 327,0699}{16 \cdot 1496 - (136)^2} = \frac{55332,1032 - 44481,5064}{5440} = 1,994595 .$$

$$b' = \frac{n \sum x_i \lg y_i - \sum x_i \cdot \sum \lg y_i}{n \sum x_i^2 - (\sum x_i)^2} ;$$

$$b' = \frac{16 \cdot 327,0699 - 136 \cdot 36,9867}{16 \cdot 1496 - (136)^2} = \frac{5233,1184 - 5030,1912}{5440} = 0,03730 .$$

Находим a и b :

$$\lg a = a' = 1,994595 ; \quad a = 98,763.$$

$$\lg b = b' = 0,03730 ; \quad b = 1,090.$$

Эмпирическое уравнение роста выпуска предприятием минеральных удобрений

$$y = 98,763 \cdot 1,090^x$$

имеет аналогичный вид. Это доказывает правильность решения задачи разными способами.

4.3 Обработка функций степенной зависимости

Если имеется степенная функция

$$y = a \cdot x^b, \quad (93)$$

то путем логарифмирования по основанию 10 её преобразуем в линейную:

$$\lg y = \lg a + b \cdot \lg x. \quad (94)$$

Обозначим $\lg y = y'$; $\lg a = a'$; $\lg x = x'$. Тогда линейная функция (94) примет вид:

$$y' = a' + b \cdot x'. \quad (95)$$

Оптимальные значения параметров a' и b находим по стандартным формулам с учетом введенных обозначений:

$$a' = \frac{\sum (x'_i)^2 \cdot \sum y'_i - \sum x'_i \cdot \sum x'_i y'_i}{n \sum (x'_i)^2 - (\sum x'_i)^2};$$

$$b = \frac{n \sum x'_i y'_i - \sum x'_i \cdot \sum y'_i}{n \sum (x'_i)^2 - (\sum x'_i)^2}.$$

Рассмотрим пример.

Пример 4.4 - Растворимость безводного хлорида аммония изменяется с температурой по степенному закону $S = a \cdot T^b$, где S – растворимость соли (г) в воде массой 100 г, T – температура, К; a и b – константы.

T , К	273	283	288	293	313	333	353	373
S , г	29,4	33,3	35,2	37,2	45,8	55,2	65,6	77,3

По результатам наблюдений методом наименьших квадратов определить оптимальные значения коэффициентов a и b в эмпирической формуле $S = a \cdot T^b$.

Решение

Константы a и b целесообразнее определить способом «линеаризации», рассмотренным ранее. Для этого прологарифмируем (по основанию 10) обе части уравнения

$$S = aT^b \quad (96)$$

и получим линейную зависимость:

$$\lg S = \lg a + b \lg T. \quad (97)$$

Обозначим

$$\lg S = Y; \quad \lg a = a'; \quad \lg T = x.$$

Тогда линейная функция примет вид:

$$Y = a' + bx$$

и задача сведется к рассмотренной ранее линейной зависимости. Все необходимые вычисления представлены в таблице 13.

Таблица 13 - Исходные данные и вспомогательные расчеты при определении оптимальных значений параметров a и b методом наименьших квадратов

T_i	S_i , эксп.	$\lg T_i = x_i$	$\lg S_i$ эксп. = y_i	x_i^2	$x_i y_i$	$S_{i \text{ расч.}}$	$\lg S_{i \text{ расч.}} = Y_i$
273	29,4	2,4362	1,4683	5,9349	3,5771	29,7	1,47
283	33,3	2,4518	1,5224	6,0113	3,7326	33,2	1,52
288	35,2	2,4594	1,5465	6,0486	3,8035	35,0	1,54
293	37,2	2,4669	1,5705	6,0854	3,8743	36,9	1,57
313	45,8	2,4955	1,6609	6,2277	4,1448	45,3	1,66
333	55,2	2,5224	1,7419	6,3627	4,3938	54,9	1,74
353	65,6	2,5478	1,8169	6,4912	4,6291	65,9	1,82
373	77,3	2,5717	1,8882	6,6137	4,8559	78,1	1,89

$$\sum_{i=1}^8 y_i = 13,2156; \quad \sum_{i=1}^8 x_i = 19,9517; \quad \sum_{i=1}^8 x_i^2 = 49,7755; \quad \sum_{i=1}^8 x_i y_i = 33,0111.$$

$$a' = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^8 x_i^2 \sum_{i=1}^8 y_i - \sum_{i=1}^8 x_i \sum_{i=1}^8 x_i y_i}{8 \sum_{i=1}^8 x_i^2 - (\sum_{i=1}^8 x_i)^2} =$$

$$= \frac{49,7755 \cdot 13,2156 - 19,9517 \cdot 33,0111}{8 \cdot 49,7755 - (19,9517)^2} = -\frac{0,8145}{0,1337} = -6,0917.$$

Из уравнения $\lg a = a'$ находим a :

$$a = 10^{-6,0917} = 8,097 \cdot 10^{-7}.$$

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{8 \sum_{i=1}^8 x_i y_i - \sum_{i=1}^8 x_i \sum_{i=1}^8 y_i}{8 \sum_{i=1}^8 x_i^2 - (\sum_{i=1}^8 x_i)^2} =$$

$$= \frac{8 \cdot 33,0111 - 19,9517 \cdot 13,2156}{8 \cdot 49,7755 - (19,9517)^2} = \frac{0,4151}{0,1337} = 3,105.$$

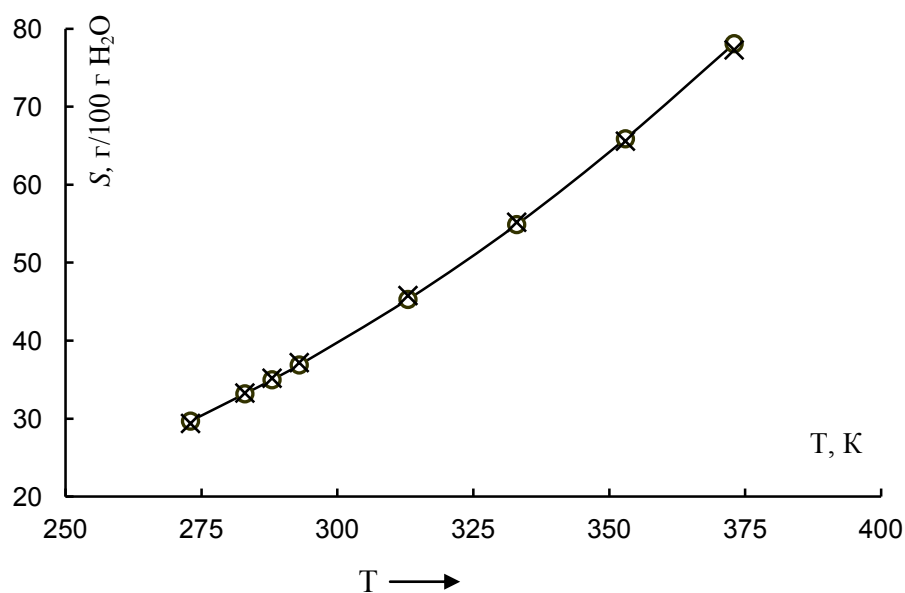
Таким образом, эмпирическое уравнение зависимости растворимости хлорида аммония от температуры имеет вид

$$S = 8,097 \cdot 10^{-7} \cdot T^{3,105}, \quad (98)$$

а в логарифмическом виде уравнение представляет линейную зависимость

$$\lg S = -6,092 + 3,105 \cdot \lg T. \quad (99)$$

Вычисленные по эмпирической формуле (99) значения растворимости хлорида аммония ($S_{i \text{ расч.}}$) для восьми значений температур представлены в таблице 13. Эти результаты можно сравнить с исходными данными ($S_{i \text{ эксп.}}$). Экспериментальные и вычисленные значения растворимости при соответствующих температурах представлены на рисунке 11 частью графика зависимости (98).



× – экспериментальные значения; о – рассчитанные значения

Рисунок 11 - Зависимость растворимости хлорида аммония от температуры

Линеаризованная зависимость (99) представлена графиком на рисунке 12.

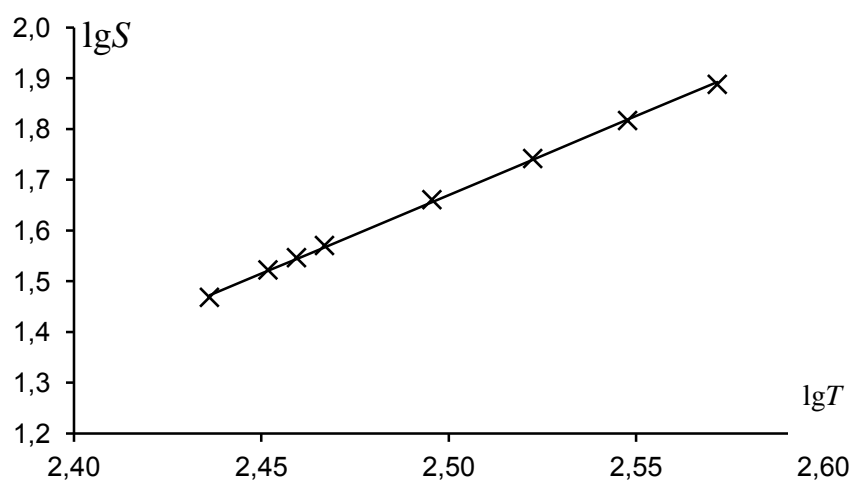


Рисунок 12 - График линеаризованной зависимости растворимости хлорида аммония от температуры, рассчитанный методом наименьших квадратов, × – экспериментальные значения

Эмпирическая формула $S = 8,097 \cdot 10^{-7} \cdot T^{3,105}$ дает возможность проводить интерполирование в пределах данных таблицы 4.5. Например, при температуре

298 К (25 °С) растворимость безводного хлорида аммония равна 39,74 г в воде массой 100 г.

Пример 4.5 - При распределении салициловой кислоты между бензолом и водой при температуре 298 К были получены следующие данные (c_1 - концентрация салициловой кислоты в водном слое (моль/дм³ или кмоль/м³); c_2 – ее концентрация в бензольном слое (моль/дм³ или кмоль/м³):

c_1 , моль/дм ³	c_2 , моль/дм ³		c_1 , моль/дм ³	c_2 , моль/дм ³		c_1 , моль/дм ³	c_2 , моль/дм ³
0,0363	0,0184		0,126	0,146		0,558	1,650
0,0668	0,0504		0,210	0,329		0,756	2,810
0,0940	0,0977		0,283	0,533		0,912	4,340

Используя закон распределения Нернста, составить эмпирическое уравнение распределения салициловой кислоты между бензолом и водой.

Решение

Согласно закону распределения Нернста, для каждой температуры отношение концентраций третьего компонента в двух равновесных фазах есть величина постоянная, не зависящая от абсолютных количеств вещества, принимающего участие в распределении.

$$\frac{c_1}{c_2} = K, \quad (100)$$

где c_1 – равновесная концентрация распределяющегося вещества в водной фазе, моль/дм³;

c_2 – равновесная концентрация распределяющегося вещества во второй (органической) фазе, моль/дм³;

K – константа распределения.

По уравнению (100) вычислим значения K :

$$K = 1,97; 1,33; 0,96; 0,86; 0,64; 0,53; 0,34; 0,27; 0,21.$$

Полученное отношение c_1/c_2 не постоянно. Это означает, что при экстракции происходит изменение молярной массы экстрагируемого вещества, вследствие ассоциации или диссоциации его молекул. В таких случаях соотношение (100) применять нельзя, а закон распределения принимает вид:

$$\frac{c_1^n}{c_2} = K, \quad (101)$$

где $n = \frac{M_2}{M_1}$ (M_1 – средняя молярная масса экстрагируемого вещества в первой (водной) фазе; M_2 – средняя молярная масса экстрагируемого вещества во второй фазе (в нашем случае в бензоле)).

Значение n показывает, во сколько раз изменяется молярная масса распределяемого вещества вследствие образования или разрушения ассоциатов (данное число n не следует отождествлять с числом параллельных определений n).

Запишем соотношение (101) в виде степенной зависимости

$$c_2 = K^{-1} \cdot c_1^n. \quad (102)$$

Таким образом, задача сводится к определению методом наименьших квадратов значений K и n степенной зависимости (102).

Способ 1

Для «выравнивания» степенной зависимости прологарифмируем (по основанию 10) обе части уравнения (102):

$$c_2 = K^{-1} \cdot c_1^n.$$

Получим: $\lg c_2 = -\lg K + n \lg c_1.$

Обозначим $\lg c_2 = Y$; $-\lg K = a$; $n=b$; $\lg c_1 = x$,

тогда получим линейную функцию $Y = a + bx$ и задача сведется к линейной.

Оптимальные значения параметров a и b находим по стандартным формулам.

Все необходимые вычисления представлены в таблице 14.

Таблица 14 - Исходные данные и вспомогательные расчеты при определении оптимальных значений параметров a и b методом наименьших квадратов

c_{1i}	c_{2i} эксп.	$\lg c_{1i} = x_i$	$\lg c_{2i} = y_i$	x_i^2	$x_i y_i$	c_{2i} расч.	$\lg c_{2i} \text{ расч.} = Y_i$
0,0363	0,0184	-1,4401	-1,7352	2,0739	2,4988	0,0184	-1,7352
0,0668	0,0504	-1,1752	-1,2976	1,3812	1,5249	0,0508	-1,2941
0,0940	0,0977	-1,0269	-1,0101	1,0545	1,0373	0,0896	-1,0477
0,1260	0,1460	-0,8996	-0,8356	0,8093	0,7517	0,1459	-0,8359
0,2100	0,3290	-0,6778	-0,4828	0,4594	0,3272	0,3411	-0,4671
0,2830	0,5330	-0,5482	-0,2733	0,3005	0,1498	0,5603	-0,2516
0,5580	1,6500	-0,2534	0,2175	0,0642	-0,0551	1,7324	0,2386
0,7560	2,8100	-0,1215	0,4487	0,0148	-0,0545	2,8705	0,4580
0,9120	4,3400	-0,0400	0,6375	0,0016	-0,0255	3,9213	0,5934

$$\sum_{i=1}^9 x_i = -6,1827; \quad \sum_{i=1}^9 y_i = -4,3309; \quad \sum_{i=1}^9 x_i^2 = 6,1594; \quad \sum_{i=1}^9 x_i y_i = 6,1546; \quad \left(\sum_{i=1}^9 x_i\right)^2 = 38,2258.$$

$$a = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2};$$

$$a = \frac{\sum_{i=1}^9 x_i^2 \cdot \sum_{i=1}^9 y_i - \sum_{i=1}^9 x_i \cdot \sum_{i=1}^9 x_i y_i}{9 \sum_{i=1}^9 x_i^2 - \left(\sum_{i=1}^9 x_i\right)^2} = \frac{6,1594 \cdot (-4,3309) - (-6,1827) \cdot 6,1546}{9 \cdot 6,1594 - 38,2258} = \frac{11,3763}{17,2088} = 0,6611; \quad ;$$

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2};$$

$$b = \frac{9 \sum_{i=1}^9 x_i y_i - \sum_{i=1}^9 x_i \sum_{i=1}^9 y_i}{9 \sum_{i=1}^9 x_i^2 - (\sum_{i=1}^9 x_i)^2} = \frac{9 \cdot 6,1546 - (-6,1827) \cdot (-4,3309)}{9 \cdot 6,1594 - 38,2258} = \frac{28,6147}{17,2088} = 1,6628.$$

Находим коэффициент распределения K и число n .

$$-\lg K = a = 0,6611, \text{ отсюда } K = 10^{-0,6611} = 0,2182.$$

$$n = b = 1,6628.$$

Значение $n = 1,6628$ показывает, что в бензольном слое салициловая кислота частично ассоциирована (преимущественно в виде димеров).

Закон распределения для данной системы выражается эмпирическим уравнением:

$$\frac{c_1^{1,6628}}{c_2} = 0,2182, \quad (103)$$

а в логарифмическом виде

$$\lg c_2 = 0,6611 + 1,6628 \cdot \lg c_1. \quad (104)$$

Значения концентрации салициловой кислоты в бензольном слое, вычисленные по формуле (104) (c_2 расч.), и $\lg c_2$ расч. приведены в последних столбцах таблицы 14.

Эмпирическая формула $\frac{c_1^{1,6628}}{c_2} = 0,2182$ дает возможность проводить

интерполирование в пределах данных таблицы 14.

Экспериментальные значения концентрации салициловой кислоты в бензольном слое в зависимости от концентрации салициловой кислоты в водном

слое отображены на рисунке 13, а по вычисленным данным ($c_{2\text{ расч.}}$) проведена линеаризованная зависимость на рисунке 14.

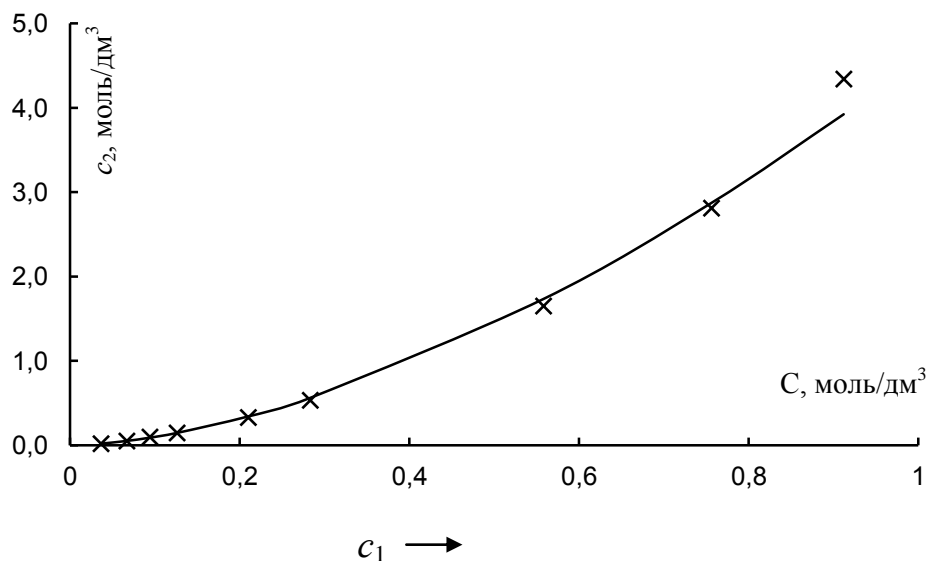


Рисунок 13 - Зависимость концентрации салициловой кислоты в бензольном слое от концентрации салициловой кислоты в водном слое, \times – экспериментальные значения

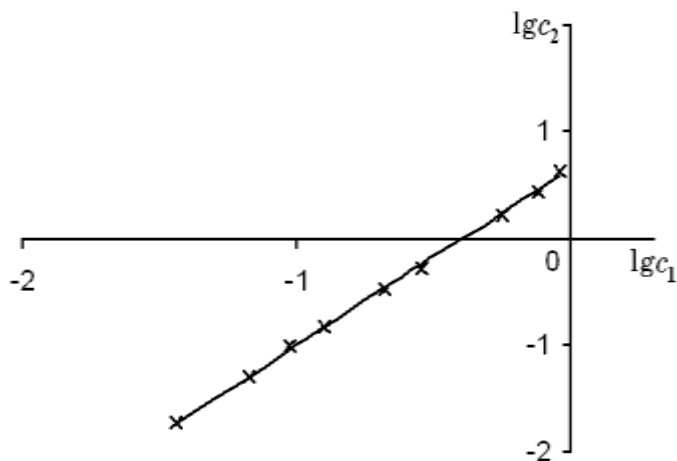


Рисунок 14 - Линеаризованная зависимость концентрации салициловой кислоты в бензольном слое от концентрации салициловой кислоты в водном слое, рассчитанная методом наименьших квадратов, \times – экспериментальные значения

Для степенной функции $y = ax^b$ в пособии не представлен вывод формул для вычисления оптимальных значений параметров a и b методом наименьших квадратов, как для показательной функции $y = a \cdot e^{bx}$. Вывод этих формул

аналогичен. Для решения задачи предпочтительней способ линеаризации степенной зависимости с последующим применением стандартных формул для нахождения оптимальных значений параметров a и b линейной зависимости.

Можно предложить и другой способ решения данной задачи. Как известно, при логарифмировании по основанию 10 степенной функции (105)

$$y = a \cdot x^b$$

получим линейную зависимость (105) вида

$$\lg y = \lg a + b \cdot \lg x ,$$

для которой нормальная система уравнений будет иметь вид:

$$\begin{cases} b \sum (\lg x_i)^2 + \lg a \sum \lg x_i = \sum \lg x_i \cdot \lg y_i \\ b \sum \lg x_i + n \cdot \lg a = \sum \lg y_i \end{cases} \quad (106)$$

Получена система двух линейных уравнений с двумя неизвестными, которая решается одним из известных способов (Гаусса или Крамера).

Решим таким способом рассмотренный ранее пример 2.

Способ 2

В соотношении $c_2 = K^{-1} c_1^n$, которое выражает распределение салициловой кислоты между бензолом и водой, обозначим:

$$c_2 = y; \quad K^{-1} = a; \quad c_1 = x; \quad n = b.$$

Получим степенную зависимость $y = a \cdot x^b$.

Исходные данные и вспомогательные расчеты для определения параметров a и b из системы уравнений (106) возьмем из таблицы 14.

$$\begin{aligned} \sum_{i=1}^9 \lg x_i &= -6,1827; & \sum_{i=1}^9 (\lg x_i)^2 &= 6,1594; \\ \sum_{i=1}^9 \lg x_i \cdot \lg y_i &= 6,1546; & \sum_{i=1}^9 y_i &= -4,3309. \end{aligned}$$

Указанные в системе (106) суммы заменяем соответствующими числами и получаем систему двух линейных уравнений с двумя неизвестными, которую решим методом Гаусса

$$\begin{cases} 6,1594b - 6,1827 \lg a = 6,1546 \\ 6,1827b + 9 \lg a = -4,3309 \end{cases} \quad (107)$$

Оставляем первое уравнение этой системы без изменения, а вместо второго записываем сумму первого и второго уравнений.

$$\begin{cases} 6,1594b - 6,1827 \lg a = 6,1546 \\ -0,0233b + 2,8173 \lg a = 1,8237 \end{cases} \quad (108)$$

Умножим первое уравнение системы (108) на 0,0233, а второе – на 6,1594 и сложим оба полученных уравнения.

$$+ \quad \begin{cases} 6,1594b - 6,1827 \lg a = 6,1546 \\ -0,0233b + 2,8173 \lg a = 1,8237 \end{cases} \begin{array}{l} 0,0233 \\ 6,1594 \end{array} \quad (109)$$

Получим уравнение с одним неизвестным ($\lg a$) :

$$6,1594b \cdot 0,0233 - 6,1827 \cdot 0,0233 \lg a - 0,0233b \cdot 6,1594 + 2,8173 \lg a \cdot 6,1594 = \\ = 0,0233 \cdot 6,1546 + 1,8237 \cdot 6,1594 .$$

$$17,3529 \lg a - 0,1441 \lg a = 11,3763 ;$$

$$17,2088 \lg a = 11,3763 ;$$

$$\lg a = 0,6611 ;$$

$$a = 10^{0,6611} = 4,582474 .$$

Подставим значение $\lg a$ во второе уравнение системы (109)

$$-6,1827b + 9 \lg a = -4,3309 .$$

Получим:

$$-6,1827b + 9 \cdot 0,6611 = -4,3309;$$

$$-6,1827b = -4,3309 - 5,9499;$$

$$-6,1827b = -10,2808;$$

$$b = \frac{10,2808}{6,1827} = 1,6628 .$$

Находим значение константы распределения K и числа n :

$$K^{-1} = a ; \text{отсюда } a = \frac{1}{K} = \frac{1}{4,582474} = 0,2182 .$$

$$n = b = 1,6628 .$$

Эмпирическое уравнение закона распределения салициловой кислоты между бензолом и водой имеет аналогичный вид

$$\frac{c_1^{1,6628}}{c_2} = 0,2182.$$

Это доказывает правильность решения задачи разными способами.

Способ решения задачи путем подстановки в нормальную систему линейных уравнений конкретных значений рационально использовать тогда, когда коэффициенты при неизвестных параметрах a и b в системе уравнений являются однозначными или двузначными числами или дробями с одним или двумя знаками после запятой. В этом случае вычисления упрощаются.

Как видно из рассмотренных примеров, при обработке многих нелинейных зависимостей методом наименьших квадратов наименее громоздким является способ линеаризации, выравнивания зависимостей с целью применения стандартных формул методов наименьших квадратов. Нелинейные зависимости, преобразованные в линейные, довольно часто встречаются в физической химии. Параметры полученных линейных зависимостей определяют важнейшие физико-химические величины, такие как изменение энтропии (ΔS), изменение энтальпии (ΔH), теплоту испарения ($\Delta H_{\text{фаз. пер.}}$) и прочие. Рассмотрим наиболее характерные примеры.

1. Зависимость константы равновесия реакции от температуры для небольшого интервала температур имеет вид:

$$\lg K_p = \frac{\Delta S}{2,30259R} - \frac{\Delta H}{2,30259R} \cdot \frac{1}{T}. \quad (110)$$

Если проведем замену переменных и констант новыми переменными и константами:

$$\lg K_p = Y; \quad \frac{\Delta S}{2,30259R} = a; \quad -\frac{\Delta H}{2,30259R} = b; \quad \frac{1}{T} = x,$$

то получим стандартную зависимость

$$Y = a + bx,$$

рассмотренную ранее, оптимальные значения параметров a и b найдем по стандартным формулам. Таким образом, оптимизация методом наименьших квадратов зависимости $\lg K_p = f(1/T)$ позволяет оценить изменение энтропии ΔS как $a \cdot 2,30259R$ и изменение энтальпии ΔH как $2,30259R \cdot b$.

2. Зависимость давления насыщенного пара (P) жидкости от температуры для небольшого интервала температур выражается уравнением Клапейрона-Клаузиуса:

$$\lg P = \text{const} - \frac{\Delta H_{\text{фаз. пер.}}}{2,30259} \cdot \frac{1}{T}. \quad (111)$$

Если ввести обозначения:

$$\lg P = Y; \quad \text{const} = a; \quad -\frac{\Delta H_{\text{фаз. пер.}}}{2,30259} = b; \quad \frac{1}{T} = x,$$

то стандартная обработка методом наименьших квадратов полученной зависимости $Y = a + bx$ позволит найти оптимальное значение теплоты испарения ($\Delta H_{\text{фаз. пер.}}$) как $b \cdot 2,30259$.

3. Константа скорости химической реакции первого порядка выражается уравнением:

$$k_c = \frac{1}{\tau} \ln \frac{c_o}{c}, \quad (112)$$

где c_o – исходная концентрация реагирующего вещества;
 c – концентрация этого вещества к моменту времени τ .

Формулу (112) можно преобразовать к виду:

$$\ln c = \ln c_o - k_c \cdot \tau \quad (113)$$

или

$$\lg c = \lg c_o - \frac{k_c}{2,30259} \cdot \tau. \quad (114)$$

Получая экспериментальным путем зависимость $c = f(\tau)$, можно методом наименьших квадратов найти оптимальное значение константы скорости k_c как $(-b) \cdot 2,30259$.

4. Уравнение изотермы адсорбции Ленгмюра

$$\Gamma = \Gamma_{\infty} \frac{kP}{1 + kP}, \quad (115)$$

может быть преобразовано к виду:

$$\frac{P}{\Gamma} = \frac{1}{k\Gamma_{\infty}} + \frac{1}{\Gamma_{\infty}} \cdot P, \quad (116)$$

где Γ – масса газа (г или мг), адсорбированного 1 г адсорбента (или 1 см² его поверхности);

k и Γ_{∞} – постоянные величины, которые зависят от природы адсорбента и от природы адсорбируемого вещества.

Если ввести обозначения

$$\frac{P}{\Gamma} = Y; \quad \frac{1}{k\Gamma_{\infty}} = a; \quad \frac{1}{\Gamma_{\infty}} = b; \quad P = x,$$

то оптимальные значения параметров a и b полученной линейной зависимости $Y = a + bx$ найдем методом наименьших квадратов по стандартным формулам. Константы k и Γ_{∞} находим следующим образом:

$$\frac{1}{\Gamma_{\infty}} = b; \quad \Gamma_{\infty} = \frac{1}{b}; \quad \frac{1}{k\Gamma_{\infty}} = a. \quad \text{Отсюда} \quad k = \frac{1}{a \cdot \Gamma_{\infty}}.$$

5. При адсорбции на границе раздела «твердое тело-жидкость» зависимость количества вещества, адсорбированного 1 г адсорбента, от равновесной концентрации вещества в растворе при постоянной температуре выражается эмпирической формулой Фрейндлиха:

$$\frac{x}{m} = a \cdot c^{1/n}, \quad (117)$$

где x – количество адсорбированного вещества, ммоль или моль;

m – масса адсорбента, г;

c – равновесная концентрация вещества в растворе, ммоль/дм³ или моль/дм³;

a и $1/n$ – постоянные величины, определяемые опытным путем.

Уравнение Фрейндлиха логарифмированием можно свести к виду:

$$\lg \frac{x}{m} = \lg a + \frac{1}{n} \lg c. \quad (118)$$

Введя обозначения

$$\lg \frac{x}{m} = Y; \quad \lg a = a'; \quad \frac{1}{n} = b; \quad \lg c = x,$$

получим линейную зависимость

$$Y = a + bx.$$

Стандартная обработка этой зависимости методом наименьших квадратов позволит найти оптимальное значение константы a как $10^{a'}$ и $1/n$ как b .

4.4 Контрольные вопросы

1. Показать применение нелинейных зависимостей в практике физико-химических исследований.

2. Дать формулировку аналитической задачи метода наименьших квадратов для квадратичной функции.

3. Показать нормальную систему уравнений для функции $SQ = \sum [(a + bx_i + cx_i^2) - y_i]^2$.

4. Описать сущность операции выравнивания (линеаризации) показательной экспоненциальной кривой.

5. Указать содержание сумм $\sum x_i$, $\sum y_i$; $\sum x_i^2$; $\sum x_i y_i$; $\sum x_i^2 y_i$; $\sum x_i^3$; $\sum x_i^4$, полученных при определении оптимальных параметров a , b и c методом наименьших квадратов для квадратичной функции вида $Y = a + bx + cx^2$.

6. Описать выравнивание показательной функции $y = a \cdot e^{bx}$; получение новой зависимости.

7. Показать сумму квадратов отклонений для показательной функции $y = a \cdot e^{bx}$ после логарифмирования ее по основанию 10.

8. Привести нормальную систему уравнений для функции $SQ = \sum (\lg a + bx_i M - y_i)^2$.

9. Привести формулы для вычисления оптимальных значений параметров a и b показательной функции $y = a \cdot e^{bx}$, выведенные методом наименьших квадратов.

10. Привести формулы для вычисления параметров a и b при сведении показательной функции $y = a \cdot e^{bx}$ логарифмированием по основанию e к линейной.

11. Записать вид нормальной системы уравнений и указать метод ее решения для показательной функции $y = a \cdot e^{bx}$ которая после логарифмирования по основанию 10 приобретает вид $\lg y = \lg a + b \cdot x \lg e$.

12. Показать линеаризацию показательной функции $y = a \cdot b^x$ (по основанию b).

13. Описать способы решения задачи методом наименьших квадратов функции $y = a \cdot b^x$ (основание b).

14. Описать способы решения системы нормальных уравнений для функции $y = a \cdot b^x$ (основание b)

15. Найти a , если: $\lg a = 0,0325$; $\lg a = 1,289$; $\lg a = 2,3251$; $\lg a = 0,2530$; $\lg a = 3,0224$; $\lg a = 1,803$.

16. Прологарифмировать степенную функцию $y = a \cdot x^b$ по основанию 10.

17. Вычислить значения параметров a' и b линейной функции $y' = a' + b \cdot x'$, полученной логарифмированием функции $y = a \cdot x^b$ по основанию 10.

18. Записать нормальную систему уравнений для степенной функции $y = a \cdot x^b$ после логарифмирования.

19. Указать необходимость изменения таблицы после приведения показательной и степенной функции к линейной зависимости при нахождении всех необходимых сумм $(\sum_{i=1}^n)$ для определения оптимальных значений параметров a и b методом наименьших квадратов.

20. Определить условие минимума для функции $y = a \cdot b^x$ (основание b).

Список использованных источников

- 1 Дёрффель, К. Статистика в аналитической химии / К. Дёрффель – М.: Мир, 1994. – 268 с.
- 2 Чарыков, А.К. Математическая обработка результатов химического анализа: учеб. пос. для вузов. / А.К. Чарыков – Л: Химия, 1984. – 168 с.
- 3 Смагунова А.Н. Методы математической статистики в аналитической химии: учеб. пособие /А.Н. Смагунова, О.М. Карпукова.-Ростов н/Д: Феникс, 2012.-346 с. ISBN 978-5-222-19507-9
- 4 Сальникова Е.В. Методы концентрирования и разделения микроэлементов: учебное пособие/ Е.В. Сальникова, Е.А. Кудрявцева.-Москва, ООО Тиру, 2012,-220 с. ISBN 978-5-93883-215-2/
- 5 Золотов, Ю.А. Основы аналитической химии: учебник для вузов: в 2 кн.- Кн.1. Общие вопросы. Методы разделения / Ю.А.Золотова [и др.]; под редакцией Ю.А. Золотова. – М.: ВШ, 1999. – 351 с.
- 5 Петерс, Д. Химическое разделение и измерение. Теория и практика аналитической химии / Д.Петерс [и др.] - М.: Химия, 1978. – 477 с.
- 6 Гмурман, В. Е. Теория вероятностей и математическая статистика: учеб. пособие / В. Е. Гмурман .- 12-е изд., перераб. - М. : Высшее образование, 2007. - 479 с. : ил.. - (Основы наук). - Прил.: с. 461-473. - Предм. указ.: с. 474-479. - ISBN 978-5-9692-0150-7
- 7 ГОСТ Р ИСО 5725-1-2002 - 5725-6-2002. ГСИ. Точность (правильность и прецизионность) методов и результатов измерений.-М.: Госстандарт России, 2002.
- 8 МИ 2336 – 2002.ГСИ. Показатели точности, правильности, прецизионности методик количественного химического анализа. Методы оценки.-Екатеринбург: УНИИМ, 2004.-45с.
- 9 Физико-химические методы анализа. Практическое руководство: учебное пособие для вузов / В.Б. Алесковский [и др.] – Л.: Химия, 1988. – 376 с.

- 10 Задачник по аналитической химии. Титриметрические методы анализа / Я.И. Коренман, П.Т. Суханов, С.П. Калинкина. – Воронеж: Воронеж. гос. технол. акад., 2001. – 336 с.
- 11 Стряпков, А.В. Математическая обработка результатов химического эксперимента: учебное пособие / А.В. Стряпков, В.А. Минаева, Т.А. Григоренко. – Оренбург: ГОУ ОГУ, 2005. –165 с.
- 12 Дорохова, Е.Н. Задачи и вопросы по аналитической химии / Е.Н.Дорохова, Г.В.Прохорова - М.: Мир, 2001. – 267 с.
- 13 Брановицкая, С.В. Вычислительная математика в химии и химической технологии / С.В. Брановицкая, Р.Б. Медведев, Ю.Я. Фиалков. – К.: Вища шк. Головное изд-во, 1986.-216 с.
- 14 Лялин, В.С. Статистика: теория и практика в Excel: учебное пособие / В.С.Лялин, И.Г. Зверева, Н.Г. Никифорова. – М.: Финансы и статистика; ИНФРА-М, 2010. – 448 с.
- 15 Харитонов, Ю. Я. Примеры и задачи по аналитической химии / Ю.Я. Харитонов, В.Ю. Григорьева - М.:ГЭОТАР - Медиа, 2008. – 304 с.
16. Гармаш, А.В. Метрологические основы аналитической химии / А.В. Гармаш, Н.М. Сорокина / МГУ, химический ф-т, каф. Аналитической химии. - Москва, МГУ, изд.3, 2012. - 47 с.
17. Боровиков, В.П. Прогнозирование в системе STATISTICA® в среде Windows. Основы теории и интенсивная практика на компьютере: учебное пособие / В.П. Боровиков, Г.И. Ивченко. – М.: Финансы и статистика, 2000. – 348 с.
18. Чекотовский, Э.В. Графический анализ статистических данных в Microsoft Excel 2000/ Э.В. Чекотовский. – М.: Издательский дом «Вильямс», 2002. – 464 с.

Приложение А

(справочное)

Функции Лапласа

Таблица А.1 - Значения функции Лапласа

$$\Theta(U) = \frac{1}{\sqrt{2\pi}} \int_0^U e^{-U^2/2} \cdot dU$$

U	$\Theta(U)$		U	$\Theta(U)$		U	$\Theta(U)$
0,01	0,0040		0,90	0,3159		1,90	0,4713
0,03	0,0120		0,95	0,3289		1,95	0,4744
0,05	0,0199		1,00	0,3413		2,00	0,4772
0,07	0,0279		1,05	0,3531		2,10	0,4821
0,10	0,0398		1,10	0,3643		2,20	0,4861
0,15	0,0596		1,15	0,3749		2,30	0,4893
0,20	0,0793		1,20	0,3849		2,40	0,4918
0,25	0,0987		1,25	0,3944		2,50	0,4938
0,30	0,1179		1,30	0,4032		2,60	0,4953
0,35	0,1368		1,35	0,4115		2,70	0,4965
0,40	0,1554		1,40	0,4192		2,80	0,4974
0,45	0,1736		1,45	0,4265		2,90	0,4981
0,50	0,1915		1,50	0,4332		3,00	0,49865
0,55	0,2088		1,55	0,4394		3,20	0,49931
0,60	0,2257		1,60	0,4452		3,40	0,49966
0,65	0,2422		1,65	0,4505		3,60	0,49984
0,70	0,2580		1,70	0,4554		3,80	0,499928
0,75	0,2734		1,75	0,4599		4,00	0,499968
0,80	0,2881		1,80	0,4641		5,00	0,499997
0,85	0,3023		1,85	0,4678			

Приложение Б
(справочное)
Значения коэффициента
Стьюдента

Таблица Б.1 - Критические значения коэффициента Стьюдента (t-критерия)
для различной доверительной вероятности p и числа степеней свободы f

f	p							
	0,80	0,90	0,95	0,98	0,99	0,995	0,998	0,999
1	3,0770	6,3130	12,7060	31,820	63,656	127,656	318,306	636,619
2	1,8850	2,9200	4,3020	6,964	9,924	14,089	22,327	31,599
3	1,6377	2,35340	3,182	4,540	5,840	7,458	10,214	12,924
4	1,5332	2,13180	2,776	3,746	4,604	5,597	7,173	8,610
5	1,4759	2,01500	2,570	3,649	4,0321	4,773	5,893	6,863
6	1,4390	1,943	2,4460	3,1420	3,7070	4,316	5,2070	5,958
7	1,4149	1,8946	2,3646	2,998	3,4995	4,2293	4,785	5,4079
8	1,3968	1,8596	2,3060	2,8965	3,3554	3,832	4,5008	5,0413
9	1,3830	1,8331	2,2622	2,8214	3,2498	3,6897	4,2968	4,780
10	1,3720	1,8125	2,2281	2,7638	3,1693	3,5814	4,1437	4,5869
11	1,363	1,795	2,201	2,718	3,105	3,496	4,024	4,437
12	1,3562	1,7823	2,1788	2,6810	3,0845	3,4284	3,929	4,178
13	1,3502	1,7709	2,1604	2,6503	3,1123	3,3725	3,852	4,220
14	1,3450	1,7613	2,1448	2,6245	2,976	3,3257	3,787	4,140
15	1,3406	1,7530	2,1314	2,6025	2,9467	3,2860	3,732	4,072
16	1,3360	1,7450	2,1190	2,5830	2,9200	3,2520	3,6860	4,0150
17	1,3334	1,7396	2,1098	2,5668	2,8982	3,2224	3,6458	3,965
18	1,3304	1,7341	2,1009	2,5514	2,8784	3,1966	3,6105	3,9216
19	1,3277	1,7291	2,0930	2,5395	2,8609	3,1737	3,5794	3,8834
20	1,3253	1,7247	2,08600	2,5280	2,8453	3,1534	3,5518	3,8495
21	1,3230	1,7200	2,2,0790	2,5170	2,8310	3,1350	3,5270	3,8190
22	1,3212	1,7117	2,0739	2,5083	2,8188	3,1188	3,5050	3,7921
23	1,3195	1,7139	2,0687	2,4999	2,8073	3,1040	3,4850	3,7676

Продолжение таблицы Б.1

24	1,3178	1,7109	2,0639	2,4922	2,7969	3,0905	3,4668	3,7454
25	1,3163	1,7081	2,0595	2,4851	2,7874	3,0782	3,4502	3,7251
26	1,315	1,705	2,059	2,478	2,778	3,0660	3,4360	3,7060
27	1,3137	1,7033	2,0518	2,4727	2,7707	3,0565	3,4210	3,6896
28	1,3125	1,7011	2,0484	2,4671	2,7633	3,0469	3,4082	3,6739
29	1,3114	1,6991	2,0452	2,4620	2,7564	3,0360	3,3962	3,8494
30	1,3104	1,6973	2,0423	2,4573	2,7500	3,0298	3,3852	3,6460
32	1,3080	1,6930	2,0360	2,4480	2,7380	3,0140	3,3650	3,6210
34	1,3070	1,6909	2,0322	2,4411	2,7284	3,9520	3,3479	3,6007
36	1,3050	1,6883	2,0281	2,4345	2,7195	9,490	3,3326	3,5821
38	1,3042	1,6860	2,0244	2,4286	2,7116	3,9808	3,3190	3,5657
40	1,303	1,6839	2,0211	2,4233	2,7045	3,9712	3,3069	3,5510
42	1,320	1,682	2,018	2,418	2,6980	2,6930	3,2960	3,5370
44	1,301	1,6802	2,0154	2,4141	2,6923	3,9555	3,2861	3,5258
46	1,300	1,6767	2,0129	2,4102	2,6870	3,9488	3,2771	3,5150
48	1,299	1,6772	2,0106	2,4056	2,6822	3,9426	3,2689	3,5051
50	1,298	1,6759	2,0086	2,4033	2,6778	3,9370	3,2614	3,4060
55	1,2997	1,673	2,0040	2,3960	2,6680	2,9240	3,2560	3,4760
60	1,2958	1,6706	2,0003	2,3901	2,6603	3,9146	3,2317	3,4602
65	1,2947	1,6686	1,997	2,3851	2,6536	3,9060	3,2204	3,4466
70	1,2938	1,6689	1,9944	2,3808	2,6479	3,8987	3,2108	3,4350
80	1,2820	1,6640	1,9900	2,3730	2,6380	2,8870	3,1950	3,4160
90	1,2910	1,6620	1,9867	2,3885	2,6316	2,8779	3,1833	3,4019
100	1,2901	1,6602	1,9840	2,3642	2,6259	2,8707	3,1737	3,3905
120	1,2888	1,6577	1,9719	2,3578	2,6174	2,8598	3,1595	3,3735
150	1,2872	1,6551	1,9759	2,3515	2,6090	2,8482	3,1455	3,3566
200	1,2858	1,6525	1,9719	2,3451	2,6006	2,8385	3,1315	3,3398
250	1,2849	1,6510	1,9695	2,3414	2,5966	2,8222	3,1232	3,3299
300	1,2844	1,6499	1,9679	2,3388	2,5923	2,8279	3,1176	3,3233
400	1,2837	1,6487	1,9659	2,3357	2,5882	2,8227	3,1107	3,3150
500	1,2830	1,6470	1,9640	2,3330	2,7850	2,8190	3,1060	3,3100

Задание № 1 (Excel)

Решение системы линейных уравнений тремя способами,

Решить систему линейных уравнений:

$$\begin{cases} 6x - 8y - 7z = 10 \\ -3x + 5y + 9z = 4 \\ 2x + 1y + 12z = -6 \end{cases}$$

На рисунке В.1 приведены все три способа решения этой системы уравнений, Аналогичным образом следует решить эту систему во время выполнения этого задания, повторив оформление решения

А	В	С	Д	Е	Г	Н	И	К	Л	М	О	Р
Способ 1. Решение системы линейных уравнений по формулам Крамера.							Способ 2. Решение системы линейных уравнений с использованием команды Сервис - Поиск решения					
6	-8	-7	=	10	D=	-35		X	Y	Z	Левая часть	Свободные члены
-3	5	9	=	4		"(1)"		6	-8	-7	10,000001	10
2	1	12	=	-6				-3	5	9	4,000001	4
								2	1	12	-6,000000257	-6
10	-8	-7			Dx=	1088	X=Dx/D=	Корни:				
4	5	9					-31,0857	-31,0857	-30,9143	7,25714		
-6	1	12										
6	10	-7			Dy=	1082	Y=Dy/D=	Начальные значения корней =0 (ячейки: J8, K8, L8). В окне Поиск решения вводятся только параметры: Изменяемые ячейки \$J\$8:\$L\$8 и Ограничения: \$M\$4:\$M\$6=\$N\$4:\$N\$6				
-3	4	9					-30,9143					
2	-6	12										
6	-8	10			Dz=	-254	Z=Dz/D=	Способ 3. Решение с помощью дистрибутивных функций				
-3	5	4					7,257143	Уравнение (1) в векторном виде DX=b. Тогда X=D ⁻¹ b, где D ⁻¹ - обратная матрица D. X, b - векторы-столбцы:				
2	1	-6						X = $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ b = $\begin{pmatrix} 10 \\ 4 \\ -6 \end{pmatrix}$				
					Проверка	10	-31,0857					
						4	-30,9143					
						-6	7,25714					

Рисунок В.1 – Решение системы уравнений методом Крамера

1 способ. Решение по формулам Крамера,

Ввод текстовой информации и его оформление не должно вызвать затруднений.

Технология получения решения и запись уравнений следующие:

1. Для каждого коэффициента уравнения, записанного в ячейки A3:C5 задайте **пользовательский формат числа**, Для 6 этот формат задается командой: [Ctrl+1] | (ДО):Формат ячеек | (Вкл):Число | Числовые форматы: (все

форматы) | Тип: 0 «х». То есть, сначала выбирается 0, а после нуля в двойных кавычках вводится **х**, Аналогично задается пользовательский формат для коэффициентов -8 , -7 первого уравнения, Только для -8 тип будет равен 0 «у», а для -7 : 0 «z =», Для всех положительных коэффициентов при z тип будет такой: для 9: +0 «z =», для 12: +0 «z =», Какой тип у 5 и 1, легко догадаться.

2. В ячейке F3 по формуле =МОПРЕД(A3:C5) рассчитайте определитель D и такое же имя присвойте этой ячейке.

3. В ячейке F7 по формуле =МОПРЕД(A7:C9) рассчитайте определитель D_x и такое же имя присвойте этой ячейке, D_x получается заменой первого столбца в матрице D на столбец свободных членов.

4. Аналогично п. 3 рассчитайте определители D_y и D_z и результирующим ячейкам присвойте имена определителей.

5. Искомые величины X, Y, Z в столбце H определите по формулам, приведенным слева от них.

6. В ячейках G18, G19, G20 делаем проверку, умножая найденные величины X, Y, Z на их коэффициенты в каждом уравнении.

2 способ. Решение системы уравнений с помощью команды *Сервис / Поиск решения*.

1. Ввести в массив **J4:L6** исходные данные – коэффициенты при неизвестных.

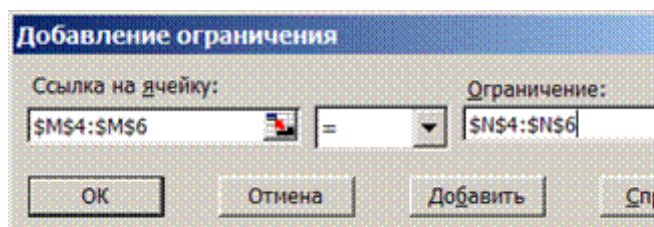
2. В массиве **M4:M6** вычисляются левые части уравнений с приближенными значениями корней уравнений, равными 0, которые задать в ячейки **J8, K8, L8**, Достаточно вычислить левую часть первого уравнения, и заполнить введенную формулу в остальные две ячейки. (Не забыть ячейки корней в формулах сделать абсолютными!).

3. В массив **N3:N6** ввести свободные члены уравнений.

4. Выполнить команду **Сервис / Поиск решения** и в диалоговом окне **Поиск решения** выполнить следующие настройки.

5. В текстовом поле *Изменяя ячейки* ввести ячейки **\$J\$8:\$L\$8**, в которых заданы нулевые значения корней.

6. Нажав кнопку **Добавить**, введите в поле *Ссылка на ячейку* выражение: **\$M\$4:\$M\$6**, в следующее поле – знак **равняется «=»**, в текстовое поле *Ограничения* – выражение **\$N\$4:\$N\$6** и нажмите кнопку **Выполнить**,



7. После поиска решения появляется окно **Результаты поиска решения**, в котором нажать кнопку **ОК**, если переключатель установлен на опции *Сохранить найденное решение*.

3 способ. Решение системы уравнений с помощью дистрибутивных и матричных функций.

Порядок решения этим способом следующий:

1. Выделить ячейки **J18:J20**, при этом ячейка J18 будет светлой, как будто не выделенная.

2. Нажмите кнопку f_x и в Мастере функций выберите категорию: *Математические*, а в поле функции: – функцию МУМНОЖ.

3. В бланке функции МУМНОЖ в поле **Массив1** установите курсор, откройте список из 10 формул (слева от строки формул) и выберите в списке последнюю строчку **Другие функции...**

4. В появившемся окне *Мастер функций* выберите функцию МОБР и в бланке этой функции в текстовое поле *Массив* задайте массив **A3:C5** – коэффициенты при неизвестных, то есть, матрицу 3х3.

5. Не нажимайте на кнопку ОК (!), а в *Строке формул* щелкните ЛКМ на название функции МУМНОЖ, Произойдет возврат в бланк функции МУМНОЖ.

6. В бланке функции МУМНОЖ установите курсор в текстовое поле **Массив2** и введите в него массив **D3:D5**.

7. Нажав клавиши **Shift + Ctrl**, щелкните на клавишу **Enter**, В ячейках **J18:J20** появятся значения вычисленных корней x , y , z .

8. Сделайте проверку решения, поместив результаты в ячейки **G18:G20**.

Приложение Г
(справочное)
Значения Q-критерия

Таблица Г.1 - Критические значения Q-критерия для различной доверительной вероятности p и числа измерений n

n	p		
	$0,90$	$0,95$	$0,99$
3	0,89	0,94	0,99
4	0,68	0,77	0,89
5	0,56	0,64	0,76
6	0,48	0,56	0,70
7	0,43	0,51	0,64
8	0,48	0,55	0,68
9	0,44	0,51	0,64
10	0,41	0,48	0,60

Приложение Д
(справочное)
Таблица значений
критерия Фишера (F-
критерия)

Таблица Д.1 - Значения критерия Фишера (F-критерия) для уровня значимости $p = 0,05$

	f_1^*										
f_2^{**}	1	2	3	4	5	6	7	8	9	10	15
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	245,95
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20

f_1^* - число степеней свободы большей дисперсии,

f_2^{**} - число степеней свободы меньшей дисперсии