

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение
высшего образования
«СЕВАСТОПОЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

ИНСТИТУТ ЯДЕРНОЙ ЭНЕРГИИ И ПРОМЫШЛЕННОСТИ

Кафедра «Химия и химические технологии»

Э.В. Ткаченко, Ю.В. Толстенко

МЕТРОЛОГИЧЕСКИЕ ОСНОВЫ ХИМИЧЕСКОГО АНАЛИЗА

Учебно-методическое пособие
для студентов очной формы обучения по направлениям
18.03.01 – Химическая технология;
05.03.06 – Экология и природопользование;
35.03.05 – Садоводство

Севастополь
СевГУ
2019

УДК [006.91:544.344.016]:004(076)

ББК 24.46я73

Т484

Р е ц е н з е н т:

Ю. А. Омельчук – кандидат химических наук,
директор Института ядерной энергии и промышленности
ФГАОУ ВО «Севастопольский государственный университет»

Ткаченко Э.В.

Т484 Метрологические основы химического анализа: учеб.-метод.
пособие / Э.В. Ткаченко, Ю.В. Толстенко. – Севастополь: СевГУ,
2019. – 47 с.: ил.

Учебно-методическое пособие составлено в соответствии с программой курсов «Лабораторный химический анализ», «Физическая химия», содержит классификацию и источники возникновения погрешностей. Представлен алгоритм оценки погрешностей прямых и косвенных измерений, рекомендации по построению графиков и их обработки. Рассмотрен метод наименьших квадратов и способы компьютерной обработки результатов измерений.

УДК [006.91:544.344.016]:004(076)

ББК 24.46я73

Учебно-методическое пособие рассмотрено и утверждено на заседании кафедры «Химия», протокол № 7 от 15 апреля 2019 г.

© Ткаченко Э.В., Толстенко Ю.В., 2019
© ФГАОУ ВО «Севастопольский
государственный университет», 2019

ВВЕДЕНИЕ

Основное назначение данного учебного пособия – познакомить обучающихся с возможностями таких компьютерных программ как Microsoft Office Excel, Origin, и показать их удобства применения для выполнения стандартных статистических расчетов в работе химика-технолога. Применение программ позволяет обеспечить весьма впечатляющую быстроту статистических расчетов, высокую надежность и достоверность результатов, возможность легкого представления данных в аналитической, графической или табличной формах.

Деятельность инженера-химика в условиях современного производства сопряжена с постоянной необходимостью проведения различных, нередко весьма сложных расчетов, обработкой экспериментальных данных в исследовательской лаборатории, обоснованием и выбором оптимальных условий проведения химического процесса. Все это приводит к широкому внедрению в исследовательскую и промышленную химию численные методы, которые позволяют реализовывать алгоритмы расчета химико-технологических расчетов. Наиболее удобным и быстрым для реализации алгоритмов технологических процессов является применение программных средств, знание и умение использовать которые является одной из обязательных этапов в процессе подготовки современного инженера-химика и химика-исследователя [1].

Прежде чем сообщить полученный тобою результат
кому бы то ни было, проверь его еще раз.
Эдмунд Ч. Беркли

1. Погрешности химического анализа. Классификация погрешностей химического анализа

Как бы старательно не было проведено количественное измерение, полученный результат, как правило, всегда отличается от истинного значения вещества, которое определяется потому, что содержит ту или иную погрешность.

Погрешностью измерения называют отклонение результата измерения от истинного значения измеряемой величины.

1.1. По способу выражения погрешности бывают абсолютные и относительные

Абсолютная погрешность

Абсолютная погрешность равна разности среднего измерения величины $\bar{X}$ и истинного значения этой величины:

$$\Delta X = \bar{X} - X_{\text{ист.}} \quad (1.1)$$

В отдельных случаях, если это необходимо, рассчитывают погрешности единичных определений:

$$\Delta X_i = X_i - X_{\text{ист.}} \quad (1.2)$$

В зависимости от того, завышает или занижает погрешность результат анализа, погрешности могут быть положительные и отрицательные.

Относительная погрешность

Относительная погрешность определяется отношением абсолютной погрешности к истинному значению величины. Относительная погрешность может быть выражена в долях или процентах и обычно знака не имеет.

$$\Delta = \frac{|\Delta X|}{X_{\text{ист.}}} \quad \text{или} \quad \Delta, \% = \frac{|\Delta X|}{X_{\text{ист.}}} \cdot 100. \quad (1.3)$$

1.2. По характеру причин, их вызывающих: систематические, случайные, промахи

К **систематическим** относят погрешности, которые *вызваны постоянно действующей причиной*, постоянны во всех измерениях. Знак систематической погрешности от одного опыта к другому не изменяется, и систематическая погрешность не влияет на результат ни в сторону увеличения, ни уменьшения.

По происхождению систематические погрешности бывают: погрешностями методическими, погрешностями приборов и реактивов, погрешностями оперативными, погрешностями индивидуальными.

Методические погрешности обусловлены определенными особенностями метода анализа, например, не количественным переходом химической реакции, частичным растворением осадка, в виде которого определяется компонент, соосаждением с осадком посторонних примесей, применением индикаторов, у которых показатель титрования pT не совпадает с pH в точке эквивалентности. Методические погрешности составляют наиболее серьезную причину получения ошибочных результатов, устранить их тяжело.

К **погрешностям, обусловленным приборами и реактивами** можно отнести погрешности, связанные с недостаточной точностью весов, использованием некалиброванных гирек, мерной посуды.

Оперативные погрешности обусловлены ошибками в работе аналитиков. Они уменьшаются с накоплением опыта работы и при повышении внимания ко всем выполняемым руками аналитическим операциям. К ним можно отнести недостаточное промывание осадка, недостаточное прокаливание осадка, неверный способ выливания растворов из пипеток. Такие погрешности трудно скорректировать.

Погрешности индивидуальные зависят от индивидуальных особенностей аналитика, например, от его способности или неспособности точно определять момент изменения окраски индикатора.

Случайными погрешностями называют неопределенные по величине и знаку погрешности, в появлении которых не наблюдается никакой закономерности. Они могут быть обусловлены, например, изменением температуры, влажности воздуха, случайными потерями вещества. В отличие от систематических погрешностей случайные погрешности не могут быть учтены и устранены введением каких-нибудь поправок. Однако они могут быть значительно уменьшены увеличением числа параллельных определений при помощи теории вероятности и математической статистики.

Промахи – это *грубые погрешности*, которые возникают при неверных измерениях, расчетах, неправильных записях. При обработке

экспериментальных данных результаты с грубыми ошибками выявляют при помощи методов математической статистики и отбрасывают.

1.3. Значащие цифры: сколько требуется цифр?

Самое слабое звено во всей цепи операции любого анализа - то, которое выполняется с наименьшей точностью. Бессмысленно стремиться проводить другие измерения, с большей точностью, чем лимитирующее.

Результаты анализа принято выражать только значащими цифрами. **Значащими** называют все достоверно известные цифры плюс первая из недостоверных. Все результаты следует округлять до первой недостоверной цифры. Для оценки недостоверности результатов физико-химических измерений необходимо учитывать реальные возможности используемых метода или методики. В качестве статистических критериев может служить, например, стандартное отклонение или доверительный интервал. В случае отсутствия таких сведений недостоверность принимают равной ± 1 в последней значащей цифре.

Если за первой недостоверной цифрой следует цифра 5, округление проводят в сторону ближайшего четного числа. Например, число 10,245 следует округлить до 10,24, а число 10,255 - до 10,26. Рекомендуются округлять конечный результат после выполнения всех арифметических действий.

Обращение с нулями

Особенно внимательными надо быть при отнесении нулей к значащим или незначащим цифрам. Нули, стоящие между другими цифрами, например, 0,104; 0,201 М; 3,050 мл; $1,04 \cdot 10^4$ являются значащими цифрами. Нули, поставленные в начале числа, например, 0,20 М; 0,2153 М; 0,201 М не являются значащими, т.к. они показывают только место запятой в десятичной дроби. Например, массу образца можно выразить в граммах или в миллиграммах (0,2205 г или 220,5 мг), при этом недостоверность массы и количество значащих цифр остаются неизменными. Недостоверность последней значащей цифры в выражении массы по-прежнему составляет ± 1 (0,0002 г или 0,2 мг).

Нули в конце числа, например, 12,270 г; 0,20 М; 0,010 М; 1,00 мл обычно считаются значащими. Если такой нуль не является значащим и используется только для указания места запятой в десятичной дроби, его следует исключить и записать число, применяя умножение на 10^n .

Например, в числе 11000, записанном в такой форме, значащими считаются все пять цифр. Если же считать, что число значащих цифр только три, то число следует записать в виде $1,10 \cdot 10^4$.

Абсолютная недостоверность числа в первом случае считается равной +1, а во втором случае она оставляет $+0,01 \cdot 10^4$ или + 100.

Сложение и вычитание

Значимость суммы или разности определяется значимостью числа с наименьшим количеством десятичных знаков. Например, при сложении чисел 28,3; 5 и 0,46 значимость определяется недостоверностью числа 5, и, следовательно, сумму чисел 33, 76 следует округлить до 34.

Числа, содержащие степени, преобразуют, приводя показатели степеней слагаемых к наибольшему. Например, при сложении чисел $2 \cdot 10^{-4}$, $6,00 \cdot 10^{-2}$ и $2,5 \cdot 10^{-3}$ их надо предоставить следующим образом: $0,02 \cdot 10^{-2}$, $6,00 \cdot 10^{-2}$ и $0,25 \cdot 10^{-2}$. Используя правило значимости суммы чисел, получаем $6,27 \cdot 10^{-2}$, поскольку значимость суммы определяется значимостью числа $6,00 \cdot 10^{-2}$, имеющего наименьшее число десятичных знаков.

Умножение и деление

Для оценки значимости произведения обычно пользуются следующим правилом: значимость произведения определяется значимостью сомножителя с наименьшим числом значащих цифр. Например, перемножение чисел 2,7 и 3,45 дает произведение, содержащее две значащие цифры - 9,3.

Извлечение квадратного корня

Относительная недостоверность результата извлечения корня вдвое меньше относительной недостоверности подкоренного числа, поэтому в некоторых случаях после извлечения корня число значащих цифр увеличивается. Например, $\sqrt{2,0} = 1,414$; относительные недостоверности числа 2,00 и результата извлечения корня равны $1 \cdot 10^{-2}$ и 0,005, соответственно, следовательно, неопределенность заключена в третьем знаке после запятой.

Возведение в степень

При возведении числа в степень относительная недостоверность результата увеличивается в число раз, равное степени. Например, при возведении в квадрат она удваивается.

Извлечение квадратного корня

Относительная недостоверность результата извлечения корня вдвое меньше относительной недостоверности подкоренного числа, поэтому в некоторых случаях после извлечения корня число значащих чисел увеличивается. Например, так как относительная недостоверность числа 1,00 равна $1 \cdot 10^{-2}$, а результат извлечения корня 0,005, т.е. неопределенность заключена в третьем знаке после запятой.

Логарифмирование

При логарифмировании число значащих чисел в мантиссе равно числу цифр, которое содержал нестепенной член числа. Характеристика логарифма не входит в число значащих цифр, так как они указывают лишь на порядок логарифмируемого числа. Например,

$$\begin{aligned}\lg 0,1 \cdot 10^{-2} &= -3,0; \\ \lg 0,10 \cdot 10^{-2} &= -3,00; \\ \lg 0,1 &= -1,0.\end{aligned}$$

Абсолютная недостоверность логарифма приблизительно в 2,5 раза меньше относительной недостоверности числа под логарифмом. Например, если логарифм известен с точностью $1 \cdot 10^{-3}$, относительная погрешность логарифмируемой величины не меньше чем $2,5 \cdot 10^{-3}$. При вычислении антилогарифма число значащих цифр уменьшается. Например, $\text{antlg} 10,23 = 1,7 \cdot 10^{10}$.

Округление чисел

При проведении любых расчетов необходимо грамотно провести округление результата, т.е. определить число значащих цифр в числе, полученном в результате арифметических действий с числами, полученными экспериментально, расчетным путем или взятыми из таблицы.

Более строгий подход основан на сравнении относительных недостоверностей сомножителей и произведения (или частного). Относительная недостоверность равна отношению абсолютной недостоверности числа к самому числу. Относительная недостоверность произведения (или частного) равна сумме недостоверностей сомножителей. Например, нужно найти частное 98:87,25. Относительные недостоверности составляют (приближенно):

$$1:98 = 1 \cdot 10^{-2} \text{ и } 0,01:87,25 = 1 \cdot 10^{-4}.$$

Следовательно, относительная недостоверность частного $0,01 + 0,0001 = 1 \cdot 10^{-2}$. При делении чисел с помощью калькулятора получаем число 1,1232. Поскольку недостоверна вторая цифра после запятой, частное следует округлить до 1,12.

Точные и приближенные вычисления

Все вычисления в количественном анализе делятся на точные и приближенные. При точных расчетах для получения окончательного результата анализа следует помнить, что точность результата определяется точностью метода, точностью мерной посуды, тщательностью измерений. Точность окончательного результата не может быть повышена путем арифметических действий над результатами измерений. Например, погрешность аккуратных определений титриметрическими методами анализа обычно колеблется в пределах 0,1 - 0,1% (отн.) Поэтому при точных экспериментальных измерениях масс на аналитических весах и объемов по бюреткам результаты прямых измерений должны содержать по четыре значащие цифры. Следовательно, все используемые в последующих расчетах величины и конечный результат также должны содержать четыре значащие цифры. Например, правильно записанная концентрация должна содержать четыре значащие цифры: правильно 0,2021 моль-л, но не 0,202 и не 0,20214; правильно записанный результат вычислений массы: $m = 0,1545$ г, но не 0,154 и не 0,15454 г.

При расчетах начальных навесок или объемов жидких проб на стадии приготовления раствора титранта или титрируемого вещества примерно заданной концентрации не требуется большая точность вычислений. В результате таких приближенных вычислений ограничиваются двумя-тремя значащими цифрами.

При всех измерениях и написании значащих цифр необходимо помнить следующее:

1) Масса взвешиваемого на аналитических весах предмета обычно измеряется с погрешностью 0,0002 г (0,2 мг). Поэтому правильно записанное число при измерении в граммах должно содержать четыре цифры после запятой, а в мг - одну цифру после запятой. Например,

8,9730 г, но не 8,973 г;

256,0 мг, но не 256 мг.

2) В титриметрических методах при работе с бюретками, позволяющими производить измерение титранта с относительной погрешностью 0,1%, объем должен быть выражен числом, содержащим четыре значащие цифры. При использовании бюреток вместимостью 5, 25, 50 мл с ценой наименьшего деления соответственно 0,01; 0,05 и 0,1 мл объем должен выражаться в первом случае числом, содержащим после запятой три цифры (например, правильно - 4,570 мл; 4,855 мл; неправильно - 4,57; 4,86), а в двух других - две цифры после запятой (например, правильно - 13,45 мл; неправильно - 13,5 мл).

3) Погрешность определений титриметрическими методами анализа при аккуратной работе колеблется в десятых долях процента. Поэтому правильно записанная концентрация должна содержать четыре значащие цифры. Например, правильно - 0,2070 моль/л, неправильно - 0,207 моль/л.

4) Результат анализа (в г) должен содержать четыре значащие цифры. Например, правильно - 0,2350 г или 235,0 мг, неправильно - 0,235 г; 0,23507 г или 235 мг; 235,07 мг.

Примеры решения задач

Задача 1. При приготовлении аммиачного буферного раствора были смешаны 10,2 мл раствора хлорида аммония ($C(\text{NH}_4\text{Cl}) = 1\text{M}$) и 2,53 мл раствора аммиака ($C(\text{NH}_3) = 2\text{M}$) и 50 мл воды. Рассчитайте объем конечного раствора, пренебрегая возможными изменениями объемов растворов при смешивании.

Решение

Складываем объемы всех растворов:

$$50 + 10,2 + 2,53 = 62,73 \text{ (мл)}$$

Наименьшее число значащих цифр после запятой имеет число 50, поэтому округляем полученную сумму до целого числа: $V = 63 \text{ мл}$.

Задача 2. Какова концентрация сульфат-ионов в растворе, полученном при смешивании равных объемов, содержащих $2 \cdot 10^{-5} \text{ M}$ сульфата натрия, $0,33 \cdot 10^{-4} \text{ M}$ сульфата калия и $5 \cdot 10^{-6} \text{ M}$ серной кислоты? Считать изменения объемов растворов при смешивании не происходит.

Решение

Сначала преобразуем числа так, чтобы уравнивать показатели степеней, приводя их к наибольшему:

$$\begin{aligned}2 \cdot 10^{-5} &= 0,2 \cdot 10^{-4} \\0,33 \cdot 10^{-4} &= 0,33 \cdot 10^{-4} \\5 \cdot 10^{-6} &= 0,050 \cdot 10^{-4}\end{aligned}$$

Складываем числа $0,2 \cdot 10^{-4} + 0,33 \cdot 10^{-4} + 0,050 \cdot 10^{-4} = 0,580 \cdot 10^{-4}$.

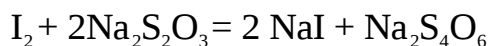
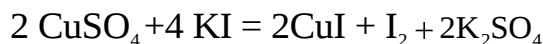
Число значащих цифр суммы должно определяться количеством их в числе $0,2 \cdot 10^{-4}$, имеющем наименьшее число значащих десятичных знаков. Поэтому округляем полученную сумму до первой цифры после запятой. Т.е. до $0,6 \cdot 10^{-4}$. В конечном объеме концентрация сульфат-ионов составляет

$$C_{\text{SO}_4^{2-}} = \frac{0,6 \cdot 10^{-4}}{3} = 0,2 \cdot 10^{-4} \text{ моль/л}$$

Задача 3. При определении содержания меди в сплаве иодометрическим методом на титрование аликвоты было израсходовано 10,00 мл раствора, полученного растворением навески 0,2000 г в объеме 100,0 мл, израсходовано 8,53 мл 0,0100 М раствора тиосульфата натрия. Вычислите содержание меди в сплаве. Молекулярная масса меди равна 63,54 г/моль.

Решение

Процесс определения катионов Cu^{2+} иодометрическим методом выражается уравнением:



Содержание меди рассчитывается по формуле

$$\omega = \frac{C_{\text{Na}_2\text{S}_2\text{O}_3} \cdot V_{p-pa(\text{Na}_2\text{S}_2\text{O}_3)} \cdot M_{\text{э}(\text{Cu}^{2+})} \cdot V_{\kappa}}{V_{\text{ал.}} \cdot 1000 \cdot m_{\text{нав.}}} \cdot 100\%$$

$$\omega = \frac{0,0100 \cdot 8,53 \cdot 63,54 \cdot 100,0}{10,00 \cdot 1000 \cdot 0,2000} \cdot 100\% = 27,095\%$$

Наименьшее число значащих цифр содержится в числе 8,53, поэтому округляем результат до сотых долей процента: 27,10 %.

1.4. Правильность. Методы проверки правильности

Систематическая погрешность определяет важнейшее понятие – **правильность**.

Под правильностью результата понимают качество измерений, которое отображает близкое значение к нулю систематической погрешности, т.е. близость среднего результата $\bar{x}$ к истинному значению $x_{ист.}$.

Правильность анализа нельзя установить посредством статистической обработки результатов. Необходимо применять экспериментальные методы проверки правильности.

Методы проверки правильности

1. Проведение анализа двумя или несколькими независимыми методами. Например, свинец можно определить гравиметрическим и титриметрическим методами. Если все методы дают одинаковые результаты, то можно быть уверенным, что анализ выполнен правильно.

2. Применение стандартных образцов.

Стандартные образцы – это эталоны химического состава материалов. Существуют организации, которые специально занимаются приготовлением и аттестацией стандартных образцов.

Аттестация стандартных образцов - это установление точного содержания отдельных элементов в них посредством анализа несколькими самыми надежными методами. Каждый стандартный образец снабжен официальным *документом-паспортом*, в котором указано аттестованное содержание отдельных элементов. Считается, что использованный метод анализа дает правильные результаты, если найденное содержание отдельных элементов в стандартном образце соответствует паспортным данным, установленным при аттестации этого образца.

3. Метод добавок.

Ведут несколько параллельных определений, прибавляя к некоторым пробам точно известные количества определяемого элемента. При правильном анализе количество элемента, найденное в пробах с добавками и без них, должно соответствовать количеству добавки.



Рисунок 1 – Правильность химического анализа

1.5. Воспроизводимость. Методы проверки воспроизводимости

Случайная погрешность определяет важнейшее понятие – воспроизводимость.

Под воспроизводимостью результатов понимают качество измерений, которое отображает близость параллельно полученных результатов одного к другому или рассеяние единичных результатов относительно среднего.

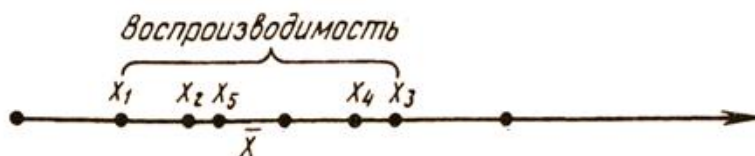


Рисунок 2 – Воспроизводимость химического анализа

О воспроизводимости результатов можно судить по следующим критериям: *отклонение от среднего d_i , среднее отклонение от среднего $\bar{d}$, дисперсия, стандартное отклонение, относительное стандартное отклонение*. Отклонения могут быть выражены как абсолютными, так и относительными величинами.

Результаты химического анализа, как и присущие этим результатам погрешности, можно рассматривать в качестве случайных. Свойства случайных величин описываются законами математической статистики. Выборка результатов анализа характеризуется определенной вероятностью p и объемом n .

Доверительная вероятность (p) – это вероятность появления результата X_i в указанной области значений.

Доверительная вероятность (p) – это число, показывающее, сколько результатов из 100 параллельных измерений попадает в интервал.

Доверительный интервал – это интервал погрешностей, соответствующий принятой доверительной вероятности P .

Из математической статистики известно, что наиболее вероятным значением измеряемой величины является математическое ожидание.

Для серии n измерений математическое ожидание $M(x)$ определяется:

$$M(x) = \sum_{i=1}^n P_i \cdot X_i, \quad (1.4)$$

где x_i – результат i -го измерения; P_i – его вероятность.

В случае равнозначных измерений

$$P_1 = P_2 = \dots P_n = \frac{1}{n} . \quad (1.5)$$

Тогда математическое ожидание

$$M(x) = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X} . \quad (1.6)$$

Таким образом, в случае равноточных измерений математическое ожидание совпадает с понятием среднего арифметического. Это правило относится к гипотетической *генеральной совокупности*, т.е. совокупности всех наблюдений, мыслимых при данных условиях. В аналитической химии используют *выборочную совокупность* или *случайную выборку*. Методами статистического анализа можно по результатам случайной выборки оценить параметры генеральной совокупности и найти наиболее вероятное значение содержания компонентов в пробе.

Распределение отклонений в выборочной совокупности отличается от нормального распределения: небольшие отклонения появляются реже, большие – чаще. Такое распределение отклонений называют *t-распределением* или *распределением Стьюдента* (статистика малых выборок). С увеличением числа параллельных определений *t-распределение* все больше приближается к нормальному распределению, а выборочное стандартное отклонение – к стандартному отклонению генеральной совокупности (при генеральной совокупности $n > 20$).

Результат единичного измерения не может быть надежной оценкой содержания компонента, который определяется в образце. Надежный результат можно получить, если провести серию параллельных измерений в идентичных условиях. Результат единичного измерения в такой серии называют *вариантой*, а вся серия – *рядом вариантов*, выборочной совокупностью или просто *выборкой*. Сумму вариантов, поделенную на число вариантов, называют *средним арифметическим выборки*:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots X_n}{n} . \quad (1.7)$$

Разность между отдельным результатом и средним выборки $\bar{X}$ называют *случайным отклонением* или *отклонением от среднего*:

$$d_i = x_i - \bar{X} . \quad (1.8)$$

Среднее арифметическое от единичных отклонений называется *средним отклонением*:

$$\bar{d} = \frac{x_i - \bar{x}}{n} . \quad (1.9)$$

Рассеяние случайной величины относительно среднего значения характеризуется **дисперсией** S^2 :

$$S^2 = \frac{\sum x_i - \bar{x}^2}{n-1} = \frac{\sum d_i^2}{n-1} . \quad (1.10)$$

Для характеристики рассеяния используют **стандартное отклонение** (**среднее квадратическое отклонение**):

$$S = \sqrt{\frac{\sum x_i - \bar{x}^2}{n-1}} = \frac{\sum d_i^2}{n-1} . \quad (1.11)$$

Относительное стандартное отклонение S_r

$$S_r = \frac{S}{\bar{x}} . \quad (1.12)$$

Среднее квадратическое отклонение среднего арифметического

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} . \quad (1.13)$$

Это безразмерная величина, объективный показатель **воспроизводимости**.

Все три величины – дисперсия, стандартное отклонение и относительное стандартное отклонение – характеризуют воспроизводимость результатов химического анализа, т.е. определяются случайной погрешностью.

Для оценки правильности могла бы служить разница между действительной величиной и найденным нами средним арифметическим. Однако обычно действительное значение определяемой величины неизвестно. Для оценки правильности используют доверительный интервал.

Статистическая надежность P , или Доверительный интервал выборочного среднего арифметического рассчитывают по формуле

$$\bar{x} \pm \Delta\bar{x} = \bar{x} \pm \frac{t_{p,f} \cdot S}{\sqrt{n}} . \quad (1.14)$$

где $\bar{x}$ – среднее арифметическое выборки;

$\Delta\bar{x}$ – точность погрешности среднего арифметического выборки;

S – стандартное отклонение выборки;
 $t_{p,f}$ – коэффициент распределения Стьюдента (табл. знач.);
 n – число вариантов (определений).

В аналитической химии применяют доверительную вероятность равную $P = 0,95$.

Распределение Стьюдента связывает между собой три основные характеристики выборочной совокупности - ширину доверительного интервала, соответствующую ему вероятность и объем выборочной совокупности.

Доверительный интервал характеризует воспроизводимость и в определенной степени правильность результатов химического анализа.

1.6. Точность анализа

Под точностью анализа понимают качество измерений, которое отображает близость полученного результата к его истинному значению или точность анализа отражает близость к нулю погрешностей всех видов.

1.7. Исключение промахов из выборки

Q-критерий

Прежде чем приступить к обработке данных с применением методов математической статистики, необходимо выявить грубые погрешности (промахи) и исключить их из числа выборки.

Для выявления промахов в ряде параллельных определений при небольшом числе измерений используют Q-критерий:

$$Q = \frac{x_1 - x_2}{W}, \quad (1.15)$$

где x_1 – сомнительный результат;
 x_2 – соседний с ним по величине результат;

Для определения Q-критерия полученные результаты располагают в порядке возрастания. Затем определяют размах варьирования (W), который равен разности между максимальным и минимальным значениями x в ряду параллельных результатов:

$$W = x_{\max} - x_{\min}. \quad (1.16)$$

Рассчитанную Q величину сравнивают $Q_{\text{табл.}}$.

Если рассчитанная величина $Q > Q_{\text{табл.}}$, подозрительный результат необходимо откинуть.

1.8. Сравнение двух или более средних значений

При анализе одного и того же образца разными методами или разными химиками-аналитиками часто возникает необходимость в сравнении двух или более средних значений полученных результатов. Проверку проводят при помощи статистического критерия F-распределения (F-критерия или критерия Фишера):

$$F = \frac{S_1^2}{S_2^2}, \quad (1.17)$$

где $S_1^2 > S_2^2$. Поэтому F-критерий всегда больше единицы.

Числовые значения F-критерия при вероятности появления $p = 0,95$ и разных значениях степеней свободы приведены в таблицах.

Если $F_{\text{эксп.}} > F_{\text{табл.}}$, то между дисперсиями существует большая разница и объединить в одну выборку результаты двух определений невозможно.

Если $F_{\text{эксп.}} < F_{\text{табл.}}$, то разность в отклонение не большая и результаты можно объединить.

2. Расчет уравнения линейного градуировочного графика, его метрологических характеристик и метрологических характеристик результата анализа

Химику часто приходится строить линейные зависимости, например, градуировочные прямые. Для обеспечения правильности результатов построения градуировочных зависимостей имеет решающее значение – ведь именно с их помощью находят содержания веществ в неизвестных образцах. Однако все результаты измерений характеризуются некоторой неопределенностью, и данные, используемые для построения градуировочных зависимостей, не составляют исключения. Они неизбежно влияют на разброс относительно прямой, и часто прямую в таких случаях проводят интуитивно, на глаз, просто приложив линейку так, чтобы точки были разбросаны относительно прямой более-менее равномерно. Естественно, более надежным было бы использование статистических методов, позволяющих определить наиболее вероятное расположение прямой.

Статистическими методами можно показать, что экспериментальный набор данных наилучшим образом описывает та прямая, для которой **сумма квадратов отклонений экспериментальных значений от рассчитанных минимальна**. Способ построения таких прямых называется **методом наименьших квадратов (МНК)**. Если x – заданная величина (например, концентрация), а y – измеряемая величина

(оптическая плотность при спектрофотометрических измерениях, площадь пика при хроматографических измерениях и т.д.), то отклонения рассчитывают вдоль вертикальной оси y . В этом случае предполагают, что значения независимой переменной x_i не содержат погрешностей. Обозначают экспериментальное значение y , соответствующее заданному значению x_i , как y_i , а рассчитанное значение y (т.е. лежащее на прямой) – как y_m . Очевидно, величина y_i в точности равна $y=a+bx$. Тогда, сумма квадратов отклонений S равна:

$$S = \sum (y_i - y_m)^2 = \sum (y_i - (a + bx_i))^2. \quad (2.1)$$

Наилучшей прямой является та, для которой величина S минимальна.

При использовании метода градуировочного графика в качестве способа определения неизвестной концентрации неизбежны погрешности, связанные с неточностью определения коэффициентов регрессионного уравнения.

Вычисление метрологических характеристик линейного графика

В химическом анализе чаще всего используют прямолинейные градуировочные графики, построенные для определенного *диапазона определяемых содержаний*, т.е. в области значений определяемых содержаний, предусмотренных данной методикой. Уравнение прямой можно записать в виде:

$$y = a + bx. \quad (2.2)$$

Уравнение (2.2) называют линейной регрессией. Параметр a градуировочной функции характеризует отрезок, отсекаемый на оси ординат, который соответствует неизбежному значению холостого опыта, а коэффициент регрессии b , характеризующий наклон градуировочной зависимости, представляет чувствительность метода (методики) анализа.

Значения параметров a и b вычисляются методом регрессионного анализа.

Если получено n экспериментальных точек $(x_1, Y_1); (x_2, Y_2); \dots (x_n, Y_n)$, используя постулат метода наименьших квадратов (МНК), можно найти параметры прямой a и b , наилучшим образом удовлетворяющие экспериментальным данным (суммирование ведется по i от 1 до n):

$$a = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2}, \quad (2.3)$$

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}. \quad (2.4)$$

Для градуировочной прямой, проходящей через начало координат и описываемой уравнением $y = bx$, МНК дает значение

$$b = \frac{\sum x_i y_i}{\sum x_i^2}. \quad (2.5)$$

Значения углового коэффициента **b** и свободного члена **a**, рассчитанные по методу наименьших квадратов, характеризует наиболее вероятное прохождение прямолинейной зависимости.

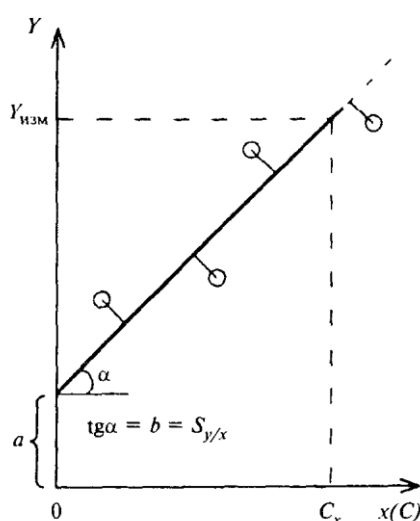


Рисунок 3 – Градуировочный график

С использованием описанных в разделе понятий можно рассчитать доверительные интервалы для параметров **a** и **b** градуировочного графика $y = a + bx$, полученного с применением МНК.

Дисперсии констант **a** и **b**

Дисперсия, характеризующая рассеяние экспериментальных значений y_i относительно рассчитанной прямой $y = a + bx$, определяется выражением:

$$S_y^2 = \frac{\sum_i^n (y_i - Y_i)^2}{n - 2}, \quad (2.6)$$

где n – число образцов сравнения, использованных для построения градуировочного графика, y_i – результат прямых измерений для i -го эталона, Y_i – результат для i -го эталона, полученный расчетным путем из уравнения $Y_i = a + bx_i$, по известному x_i и значениям a и b , найденным по приведенным выше формулам.

Величину $\sum_{i=1}^n (y_i - Y_i)^2$ можно рассчитать по формуле

$$\sum_{i=1}^n (y_i - Y_i)^2 = \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n y_i - b \sum_{i=1}^n x_i y_i, \quad (2.7)$$

где i меняется от 1 до n .

Стандартное отклонение точек от найденной зависимости

Для характеристики рассеяния результатов в выборочной совокупности рассчитывают также и стандартное отклонение, которое рассчитывается по формуле

$$S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - Y_i)^2}{n-2}}, \quad (2.8)$$

Дисперсии параметров a и b рассчитывают по формулам

$$S_a^2 = S_y^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = S_y^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2.9)$$

$$S_b^2 = S_y^2 \frac{n}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = S_y^2 \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2.10)$$

где $\bar{x}$ – среднее из всех значений x_i :

$$\bar{x} = \sum_{i=1}^n x_i / n. \quad (2.11)$$

Из дисперсий можно рассчитать стандартные отклонения и доверительные интервалы для a и b :

$$\Delta a = S_a \cdot t_{p,f}, \quad (2.12)$$

$$\Delta b = S_b \cdot t_{p,f}, \quad (2.13)$$

где p – принимают равным 0,95, а $f = n-2$.

3. Проверка коэффициента a уравнения регрессии на статистическую значимость по критерию Стьюдента

Проверка статистической значимости параметров регрессионного уравнения (коэффициентов регрессии) выполняется по t -критерию Стьюдента, который рассчитывается по формуле

$$t_{p,f} = \frac{|a|}{S_a}, \quad (2.14)$$

где a - коэффициентов регрессии; S_a - стандартное отклонение параметра.

Рассчитанное значение критерия Стьюдента сравнивают с его табличным значением при выбранной доверительной вероятности (0,95) и числе степеней свободы $n - k - 1$, где N - число точек, k - число переменных в регрессионном уравнении (например, для линейной модели $y = a + bx$ подставляем $k = 1$, тогда $f = n - 2$).

Если вычисленное значение t_p выше, чем табличное, то коэффициент регрессии является значимым с данной доверительной вероятностью. В противном случае есть основания для исключения соответствующей переменной из регрессионной модели.

Если коэффициент a статистически незначим рекомендуется его исключить из уравнения регрессии. Это не приведет к существенной потери качества модели, но сделает ее более конкретной.

$$\Delta a = S_a \cdot t_{p,f}. \quad (2.15)$$

Если при этом a будет $< \Delta a$, то значение a считают незначимым и переходят к уравнению прямой вида $y = b'x$, где b' рассчитывается по формуле (2.16).

Проверку значимости константы a можно провести еще другим способом.

4. Проверка коэффициента a уравнения регрессии на статистическую значимость по F-критерию

В ряде случаев значение a настолько мало, что этим членом уравнения прямой можно пренебречь. При зависимости $y = b'x$ вычисления значительно упрощаются. Тогда:

$$b' = \frac{\sum x_i y_i}{\sum x_i^2}. \quad (2.16)$$

$$S_y^2 = \frac{\sum_i^n (y_i - Y_i)^2}{n-1}, \quad (f = n-1) \quad (2.17)$$

$$SQ' = \sum_i^n (y_i - Y_i)^2 = S_y^2(n-1) = \sum y_i^2 - b' \sum x_i y_i \quad (2.18)$$

$$S_b^2 = S_y^2 / \sum x_i^2, \quad (f = n-1) \quad (2.19)$$

$$\Delta b = S_b \cdot t_{p,f} = t_{p,f} \sqrt{S_y^2 / \sum x_i^2} \quad (2.20)$$

В ситуации, когда значение a уравнения $y = a + vx$ мало отличается от нуля и возникает предположение, что уравнение может иметь вид $y = v'x$, осуществляют проверку константы a по следующей схеме:

1. Вычисляют сумму квадратов по уравнению (2.21):

$$SQ = \sum_i^n (y_i - Y_i)^2 = S_y^2(n-2) = \sum y_i^2 - a \sum y_i - b \sum x_i y_i, \quad (f = n-2)$$

При проведении вычислений по данной формуле необходимо выполнять при достаточно большом числе знаков после запятой, так как искомую сумму квадратов часто ищут для весьма близких значений. Поэтому на данном этапе даже незначительные погрешности в вычислении могут привести к большим погрешностям.

При этом $SQ' > SQ$.

2. Затем находят критерий Фишера по формуле

$$F = \frac{(SQ' - SQ)}{S_y^2} \quad (2.22)$$

Сравнивают значения F , если $F_{\text{эсп.}} < F_{\text{табл.}}$, то имеются основания для перехода к уравнению $y = v'x$.

Стандартные ошибки коэффициентов используются для оценивания параметров уравнения регрессии.

Стандартное отклонение точек от найденной зависимости рассчитывают по формуле:

$$S_y = \sqrt{\frac{\sum_i^n (y_i - Y_i)^2}{n-1}} = \sqrt{\frac{\sum_i^n (y_i - (b'x_i))^2}{n-1}}, \quad (2.23)$$

Стандартное отклонение рассчитанного значения наклона регрессионной зависимости рассчитывают по формуле

$$S_b = \sqrt{\frac{S^2}{\sum x_i^2}}, \quad (2.24)$$

Точность значений параметров a и b рассчитывают по уравнениям:

$$\varepsilon_a = \frac{t_{p,f} \cdot S_a}{\sqrt{n}}, \quad (2.25)$$

$$\varepsilon_b = \frac{t_{p,f} \cdot S_b}{\sqrt{n}}, \quad (2.26)$$

$t_{p,f}$ – коэффициент Стьюдента при заданной доверительной вероятности p ($p = 0,95$) и числа степеней свободы f ($f = n - 2$).

Если $a = 0$, то при определении точности значения b число степеней свободы f рассчитывается как $f = n - 1$.

5. Границы доверительного интервала для рассчитанного значения зависимой переменной

Поскольку параметры a и b , рассчитываемые по МНК, являются случайными величинами, то и значения $y_i = a + bx_i$ тоже являются случайными. Вследствие неизбежной погрешности при определении параметров a и b необходимо рассматривать y_k для одного заданного x_k так же, как случайную величину. Доверительный интервал для вычисленного значения y_k рассчитывается по формуле

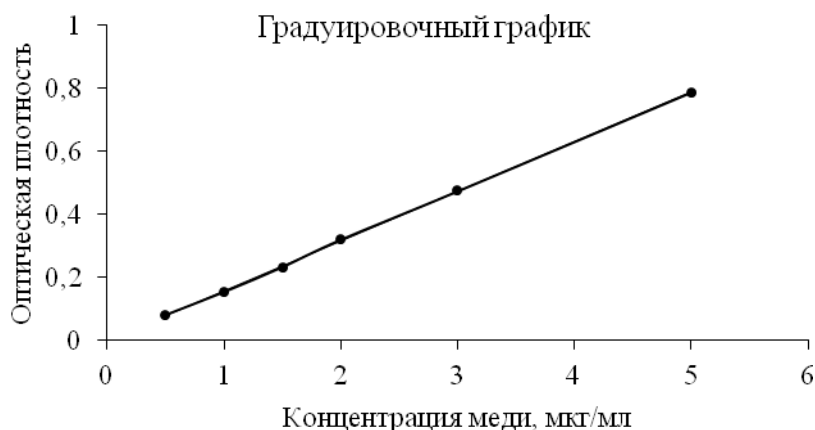
$$\Delta y_k = t_{p,f} \sqrt{S^2 \left(\frac{1}{n} + \frac{n(x_k - \bar{x})^2}{n \sum x_i^2 - (\sum x_i)^2} \right)}, \quad (2.27)$$

Доверительный интервал при этом зависит от разности $(x_k - \bar{x})$ и становится тем ближе, чем дальше лежит x_k от среднего $\bar{x}$.

Примеры решения задач

Пример 1. Для спектрофотометрического определения меди построена экспериментальная градуировочная зависимость:

х: C_{Cu} , мкг/мл	0,5	1,0	1,5	2,0	3,0	5,0
у: А	0,081	0,154	0,233	0,320	0,474	0,788



Рассчитайте параметры a и b линейной градуировочной зависимости $y=a+bx$, укажите их доверительные интервалы.

Решение

x_i	y_i	x_i^2	$x_i y_i$
0,5	0,081	0,25	0,0405
1,0	0,154	1,00	0,1540
1,5	0,233	2,25	0,3495
2,0	0,320	4,00	0,6400
3,0	0,474	9,00	1,4220
5,0	0,788	25,00	3,9400
$\sum x_i = 13$	$\sum y_i = 2,05$	$\sum x_i^2 = 41,5$	$\sum x_i y_i = 6,546$
$(\sum x_i)^2 = 169$			

1) Параметры градуировочной зависимости рассчитывают по формулам (2.1) и (2.2).

$$a = \frac{2,05 - 41,5 - 13,0 \cdot 6,546}{6 \cdot 41,5 - 169} = -0,000283$$

$$b = \frac{6 \cdot 6,546 - 13,0 \cdot 2,05}{6 \cdot 41,5 - 169} = 0,158$$

2) Для расчета доверительных интервалов найденных параметров необходимо рассчитать дисперсию S^2 , где значения Y_i вычисляются как $y_i = a + bx_i$

x_i	y_i	Y_i	$(y_i - Y_i)^2$
0,5	0,081	0,078625	$5,64 \cdot 10^{-6}$
1,0	0,154	0,157538	$1,25 \cdot 10^{-5}$
1,5	0,233	0,236450	$1,19 \cdot 10^{-5}$
2,0	0,320	0,315363	$2,15 \cdot 10^{-5}$
3,0	0,474	0,473188	$6,60 \cdot 10^{-7}$
5,0	0,788	0,788838	$7,01 \cdot 10^{-7}$
			$\sum (y_i - Y_i)^2 = 5,29 \cdot 10^{-5}$

3) Дисперсию значений y_i относительно рассчитанной прямой $y=a+bx$, находят по формуле (2.6):

$$S^2 = \frac{\sum_i^n (y_i - Y_i)^2}{n - 2},$$

$$S^2 = \frac{5,29 \cdot 10^{-5}}{6 - 2} = 1,32318 \cdot 10^{-5}$$

4) Стандартное отклонение определяется выражением:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_i^n (y_i - Y_i)^2}{n - 2}} = \sqrt{1,32 \cdot 10^{-5}} = 3,64 \cdot 10^{-3}$$

5) Зная значение S^2 , определяют дисперсии параметров a и b по формулам (2.9) и (2.10).

$$S_a^2 = S^2 \cdot \frac{\sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2} = 1,32 \cdot 10^{-5} \cdot \frac{41,5}{6 \cdot 41,5 - 169} = 6,85 \cdot 10^{-6}$$

$$S_b^2 = S^2 \cdot \frac{n}{n \sum x_i^2 - (\sum x_i)^2} = 1,32 \cdot 10^{-5} \cdot \frac{6}{6 \cdot 41,5 - 169} = 9,90 \cdot 10^{-7}$$

6) Определение стандартных отклонений параметров a и b

$$S_a = \sqrt{S_a^2} = \sqrt{6,85 \cdot 10^{-6}} = 2,62 \cdot 10^{-3}$$

$$S_b = \sqrt{S_b^2} = \sqrt{9,90 \cdot 10^{-7}} = 9,95 \cdot 10^{-4}$$

7) Доверительные интервалы величин a и b рассчитывают по формулам (2.12) и (2.13)

Используя таблицу приложения «Значения t при различной доверительной вероятности», определяют коэффициент Стьюдента ($t_{p,f}$). При $f = 4$ ($f = n - 2 = 6 - 2$) и $p = 0,95$, $t_{p,f} = 2,78$.

$$\Delta a = S_a \cdot t_{p,f} = 2,62 \cdot 10^{-3} = 7 \cdot 10^{-3}$$

$$\Delta b = S_b \cdot t_{p,f} = 9,95 \cdot 10^{-4} = 3 \cdot 10^{-3}$$

8) С учетом доверительного интервала параметры a и b равны:

$$a = (-0,3 \pm 7) \cdot 10^{-3}$$

$$b = 0,158 \pm 0,003$$

9) Проверка значимости a

$$t_{p,f} = \frac{|a|}{S_a}$$

$$f = n - 2$$

$$t_{p,f} = \frac{|-2,83 \cdot 10^{-4}|}{2,62 \cdot 10^{-3}} = 0,11$$

по таблице Стьюдента определяют критические значения $t_{p,f}$. Если $t_{p,f}$ расчетное меньше табличного, то отклонение свободного члена a от нуля можно считать случайным.

В задаче $t_{p,f} \text{ (расч.)} = 0,11$, а $t_{p,f} \text{ (табл.)} = 2,78$. Таким образом, величина свободного члена a незначимо отличается от нуля.

10) Границы доверительного интервала для рассчитанного значения зависимой переменной

$$\Delta y_k = 2,78 \sqrt{1,32 \cdot 10^{-5} \left(\frac{1}{6} + \frac{6(1,5 - 2,17)^2}{6 \cdot 41,5 - 169} \right)} = 0,0046$$

Пример 2. Используя градуировочную зависимость (пример 1), рассчитайте концентрацию меди в анализируемом растворе, укажите

доверительный интервал, если значение оптической плотности равно 0,517.

Решение

Концентрацию меди находят по уравнению градуировочной зависимости

$$x = \frac{y - a}{b} = \frac{0,517 + 0,0003}{0,158} = 3,27 \text{ мкг / мл}$$

Для расчета доверительного интервала полученной величины определяют ее стандартное отклонение по формуле:

$$S(x) = \frac{S}{b} \sqrt{\frac{1}{m} + \frac{1}{n} + \frac{S_b^2}{b^2} \frac{(\hat{y} - \bar{y})^2}{S^2}},$$

где m – число параллельных измерений сигнала анализируемого образца;
 n – число данных (образцов сравнения), использованных для градуировки; $\bar{y}$ – среднее значение сигнала анализируемого образца; $\hat{y}$ – среднее значение сигнала для всех образцов сравнения:

$$\bar{y} = \frac{\sum_{i=1}^m y_i}{n}.$$

В данном примере $m = 1$, поэтому

$$S(x) = \frac{3,64 \cdot 10^{-3}}{0,158} \sqrt{1 + \frac{1}{6} + \frac{9,90 \cdot 10^{-7}}{(0,158)^2} \frac{(0,517 - 2,05 / 6)^2}{1,32 \cdot 10^{-5}}} = 0,026.$$

Доверительный интервал рассчитывают по формуле Стьюдента, как

$$\Delta x = S_x \cdot t_{p,f} = 0,026 \cdot 2,78 = 0,07$$

$$x = (3,27 \pm 0,07) \text{ мкг / мл.}$$

6. Компьютерная обработка результатов измерений

Мощным инструментальным средством при выполнении статистических исследований является использование компьютерной техники.

6.1. Построение градуировочного графика в программе Microsoft Office Excel

При выполнении статистических исследований большой известностью пользуется приложение Microsoft Excel, которое включает в себя программную надстройку "Пакет анализа" и богатую библиотеку из большого числа статистических функций.

6.1.1. Построение линейного тренда

Для получения линейного тренда необходимо выполнить следующий алгоритм:

1. Создать и заполнить таблицу экспериментальными данными.
2. Построить градуировочный график.

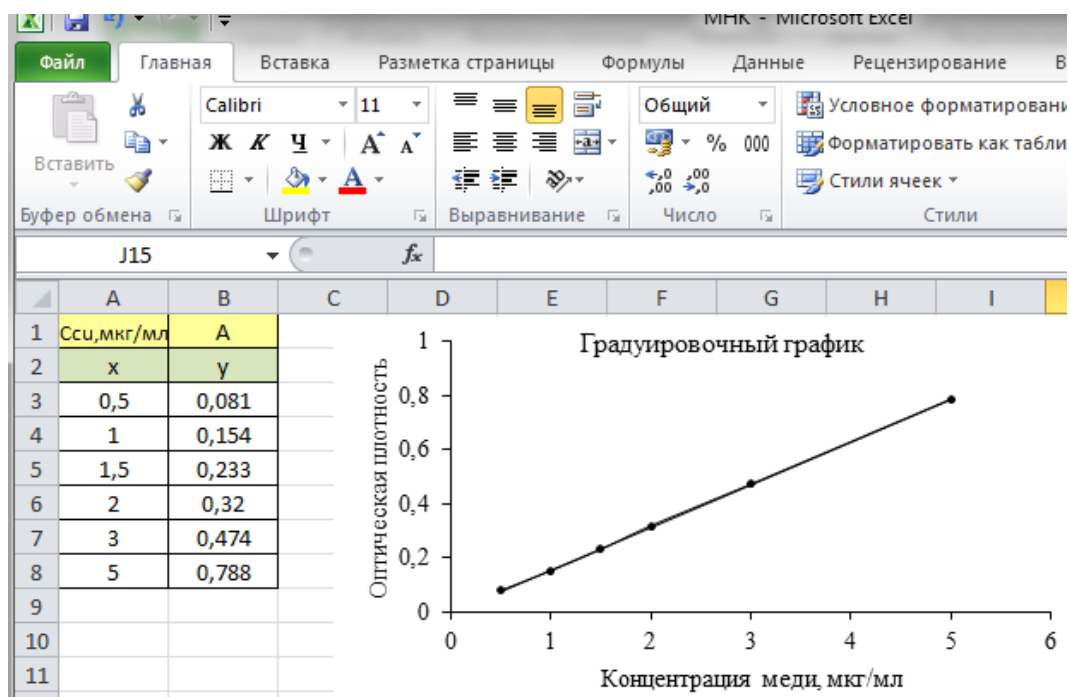


Рисунок 4 – Построение градуировочного графика

3. Щелкнуть на поле диаграммы «Градуировочный график».

4. Выполнить команду *Диаграмма* → *Добавить линию тренда*
 5. В открывшемся окне на вкладке *Тип* выбрать *Линейный тренд*
 6. Перейти на вкладку *Параметры*, установить галочки на флажках: *показывать уравнения на диаграмме и поместить на диаграмму величину достоверности аппроксимации R^2* . Щелкнуть ОК.
- Полученная диаграмма представлена на рис. 5.

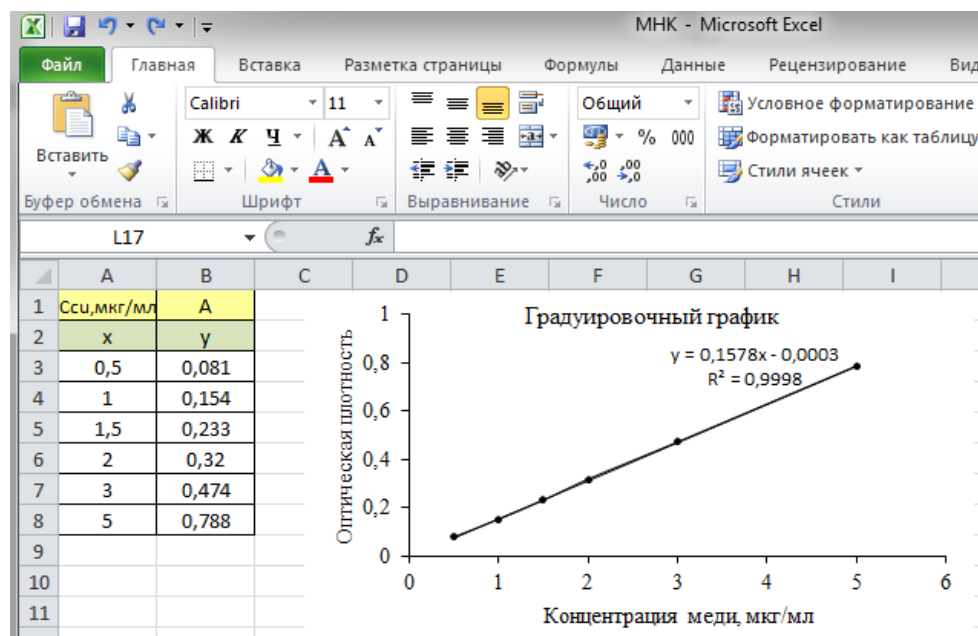


Рисунок 5 – Построение линии тренда в EXCEL для анализа графика

Расчет коэффициентов a и b без построения графика

Использование статистических функций Excel позволяет рассчитать величины углового коэффициента (a), свободного члена (b) и коэффициента детерминации линейной зависимости (R^2) без построения графика.

1. Откройте новый файл и введите данные x и y .
2. Выделите ячейку C10, нажмите f_x (вставить функцию), выберите категорию (статистические), выберите функцию «отрезок», выделите эту строку и нажмите ОК. В строке *Известные значения x* введите диапазон адресов B3:B8, а в строке *Известные значения y* C3:C8. Нажмите кнопку ОК, и величина свободного члена появится в ячейке C11.
3. Аналогично поступите с ячейкой C11, выделив ее и выбрав в качестве имени функции «наклон». Значение углового коэффициента появится в ячейке C11.
4. Выделите ячейку C12 и выберите имя функции «квпирсон». В ячейке C12 появится значение R^2 . Рассчитанные коэффициенты представлены на рис. 6.

	A	B	C	D	E	F	G	H	I	J
1	№п/п	Ccu, мкг/мл	A							
2	1	x	y							
3	2	0,5	0,081							
4	3	1	0,154							
5	4	1,5	0,233							
6	5	2	0,32							
7	6	3	0,474							
8	7	5	0,788							
9	8									
10	9	b	-0,00029	величина свободного члена						
11	10	a	0,157825	значение углового коэффициента						
12	11	R²	0,999841	коэффициент детерминации линейной зависимости						

Рисунок 6 – Аналитическое определение коэффициентов линейной зависимости

Сравните полученные значения с приведенными выше. Как видно, они совпадают.

R^2 – коэффициент детерминации линейной зависимости. График наиболее точно соответствует значениям исходных данных в действительности, когда величина R^2 близка к 1.

6.1.2. Надстройка «Пакет анализа»

Средства, которые включены в надстройку *Пакет анализа*, доступны через команду *Сервис*→*Анализ данных*. (Если команды *Анализ данных* нет в меню *Сервис*, подключите эту надстройку. Для этого выполните команду *Сервис*→*Надстройки* и в открывшемся диалоговом окне *Надстройки* в списке *Доступные надстройки* установите флажок напротив опции *Пакет анализа*). Команда *Сервис*→*Анализ данных* открывает одноимённое диалоговое окно, в списке *Инструменты анализа* которого следует выбрать необходимое средство. После выбора какого-либо средства (и последующего щелчка на кнопке *ОК*) открывается диалоговое окно этого средства.

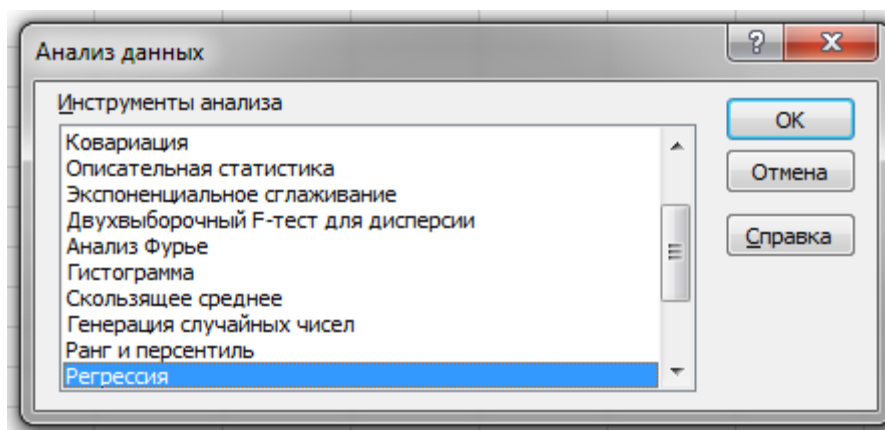
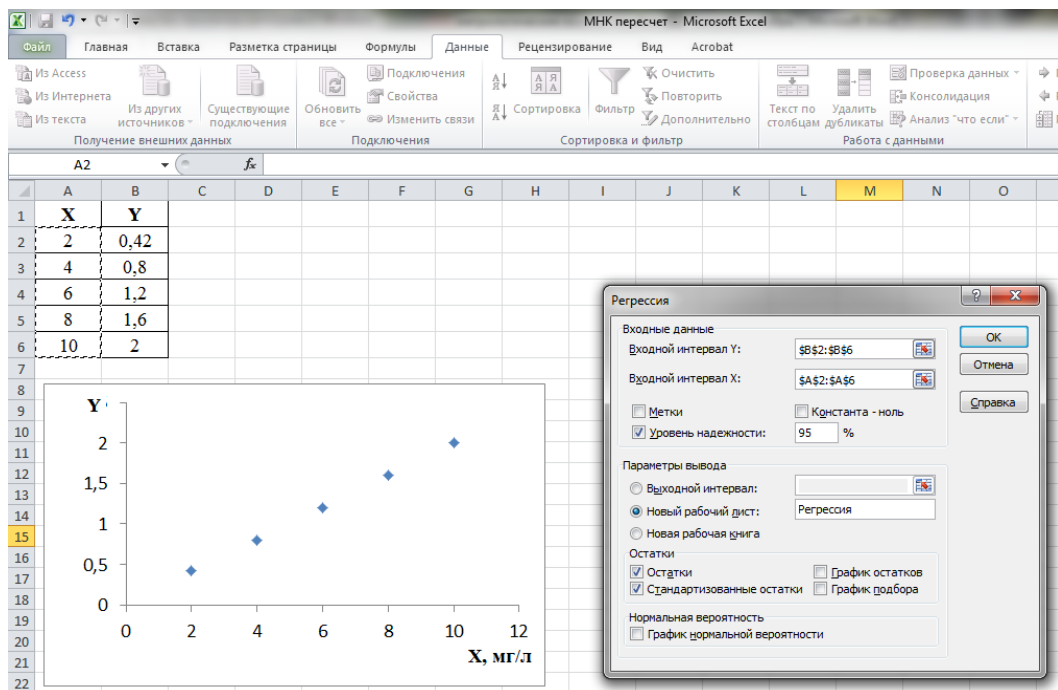


Рисунок 7 – Диалоговое окно «Регрессия» в надстройке «Пакет анализа»

В главном меню запустим серию команд Сервис→Анализ данных→Регрессия (рис.7). В появившемся диалоговом окне заполним поля ввода данных для обоих параметров Y и X ; для этого в каждое окно (Входной интервал Y и Входной интервал X) поместим значения.

Исследуется некоторая зависимость $y = f(x)$. Будем исходить из предположения, что эта зависимость описывается линейным уравнением. Об этом предварительно можно судить по виду построенного графика (рис. 8).

Отметим Уровень надёжности (доверительную вероятность), равный 95%, вывод остатков и стандартизованных остатков. Укажем в окне вывода Выходной интервал рабочий лист, где будет формироваться весь блок получаемых статистических показателей. После чего нажмём кнопку ОК. На рис. 9 представлены рассчитанные статистические показатели под заголовком «Вывод итогов». Интерес представляют показатели, которые именованы как Коэффициенты. Один из них назван Y -пересечение, а второй – Переменная $X1$. Это и есть нужные коэффициенты регрессии: свободный член a и коэффициент b при аргументе.



Регрессия - Microsoft Excel							
1	ВЫВОД ИТОГОВ						
2							
3	Регрессионная статистика						
4	Множественный R	0,999948989					
5	R-квадрат	0,99989798					
6	Нормированный R-квадрат	0,999863973					
7	Стандартная ошибка	0,007302967					
8	Наблюдения	5					
9							
10	Дисперсионный анализ						
11		df	SS	MS	F	Значимость F	
12	Регрессия	1	1,56816	1,56816	29403	4,37351E-07	
13	Остаток	3	0,00016	5,33333E-05			
14	Итого	4	1,56832				
15							
16		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
17	Y-пересечение	0,016	0,007659417	2,088931871	0,12790722	-0,008375683	0,0403757
18	Переменная X 1	0,198	0,001154701	171,4730299	4,3735E-07	0,194325228	0,2016748
19							
20							
21							
22	ВЫВОД ОСТАТКА						
23							
24	Наблюдение	Предсказанное Y	Остатки	Стандартные остатки			
25	1	0,412	0,008	1,264911064			
26	2	0,808	-0,008	-1,264911064			
27	3	1,204	-0,004	-0,632455532			
28	4	1,6	2,22045E-16	3,51083E-14			
29	5	1,996	0,004	0,632455532			
30							

Рисунок 9 – Представление результатов анализа

6.2. Построение градуировочного графика в программе OriginPro

OriginPro - это мощнейший научный пакет для анализа и обработки данных. Считается стандартом де-факто во многих научных лабораториях. OriginPro представляет собой универсальное средство обработки статистических и математических функций, обработки массивов данных, построения графиков этих функций и графиков по массивам данных, а также инструмент разработки узкоспециализированных математических программ и средств визуализации графических данных.

Построение линейной регрессии

1. При запуске Origin в рабочем пространстве открывается окно, предназначенное для ввода данных (рис.10).

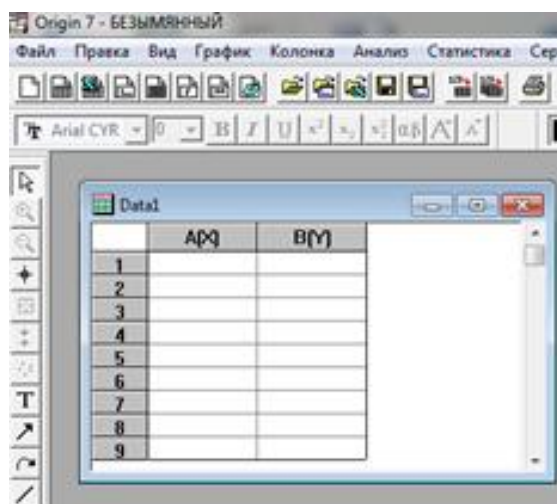


Рисунок 10 – Диалоговое окно «Ввод данных»

2. Заполните таблицу своими данными (рис. 11). Далее в Меню найдите *График→Точка→Выбор колонок для графика→Ок*

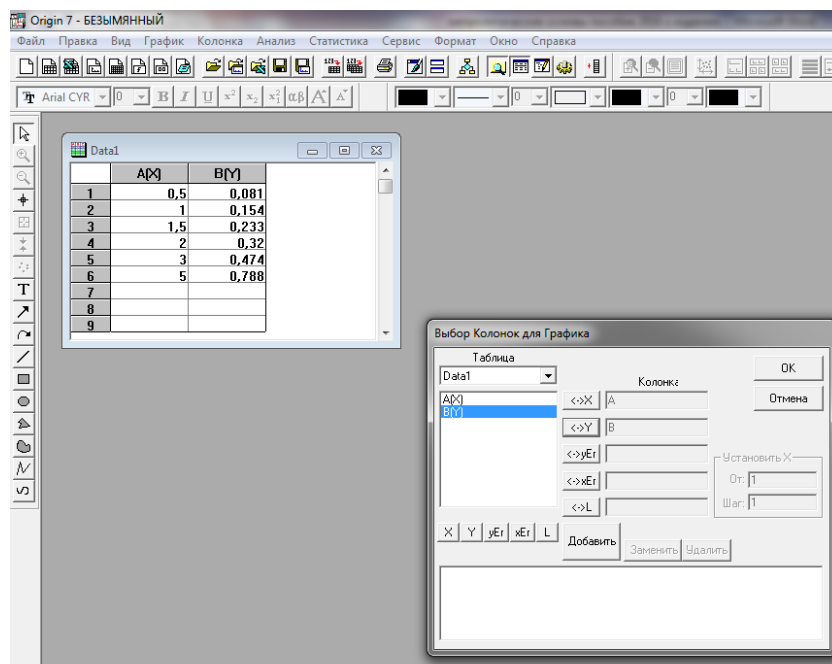


Рисунок 11 – Создание таблицы с результатами эксперимента

3. Появится окно с графиком (рис.12).

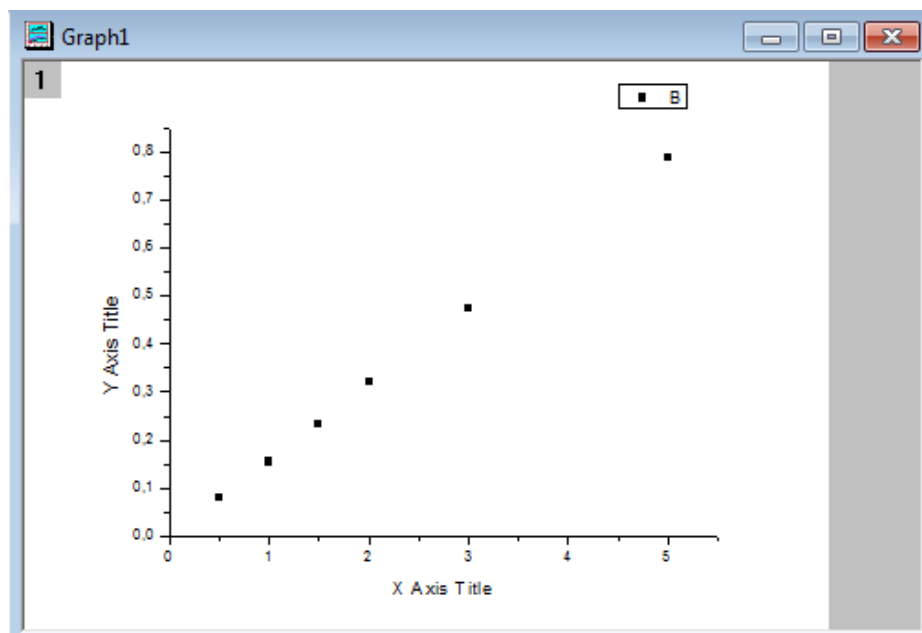


Рисунок 12 – Построение графика

4. В таком виде график представлять нельзя. Надо подписать оси x и y. Выбрать шрифт Times New Roman и размер шрифта 36 ppt (рис. 13). Когда будете сохранять в Microsoft Word, текст будет хорошо читаемым.

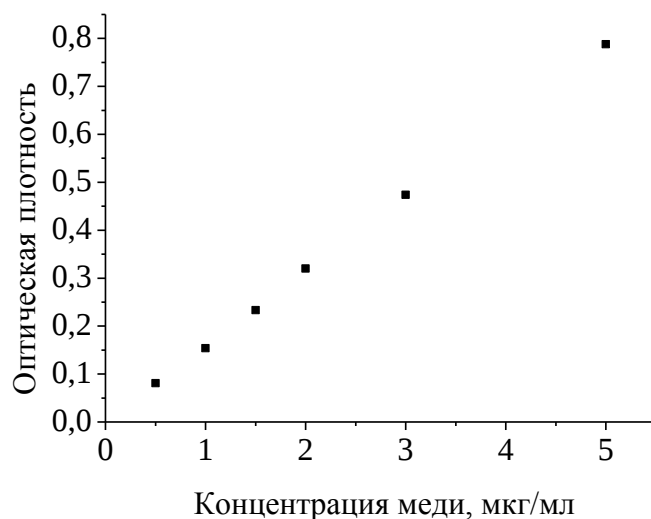


Рисунок 13 – Подпись осей

5. Для проведения регрессионного анализа откройте окно *График*. На панели управления найдите *Сервис*→*Линейная обработка*.

6. Точки на графике соединятся красной линией (линейная регрессия). В правом нижнем углу появится таблица Linear Regression for Data1_B. Данные этой таблицы скопируйте и поместите на график (как показано на рис. 14).

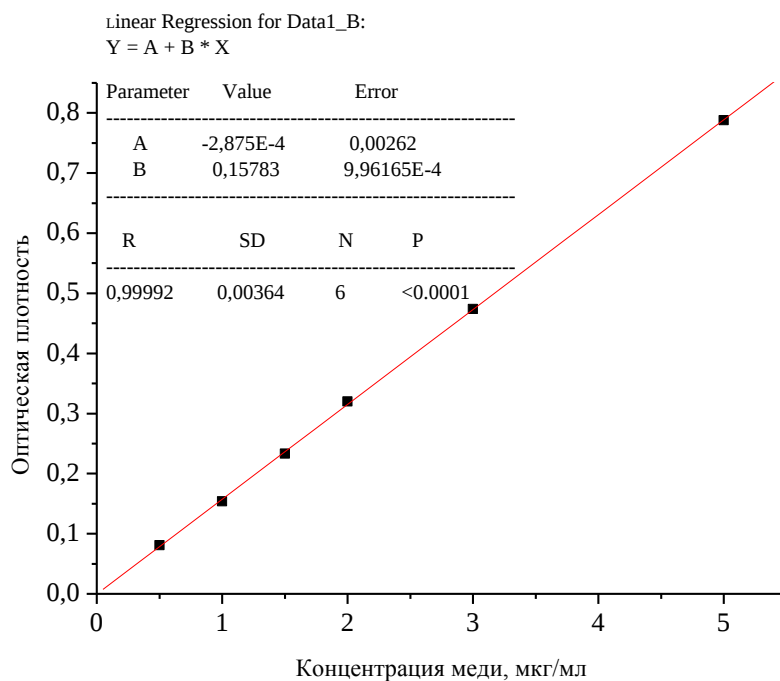


Рисунок 14 – Графическое представление результатов анализа

Обработка данных потенциометрического титрования

1. Создайте новый проект с листом данных, состоящим из двух столбцов: A(X) - объем титранта, мл; B(Y) - значения pH.
2. Выделите данные и постройте (*Plot*→*Scatter*) график экспериментальных точек (рис. 15).

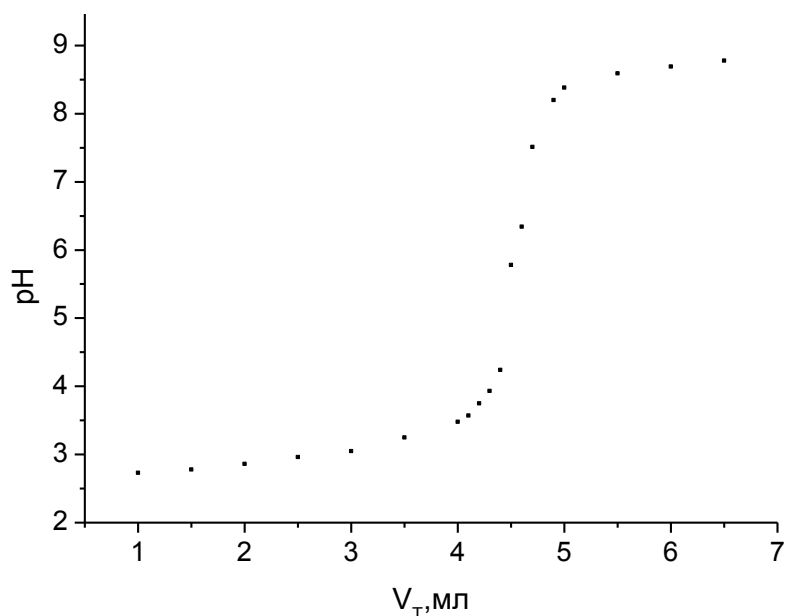


Рисунок 15 – Логарифмическая кривая титрования

3. Продифференцируйте (*Analysis* → *Calculus* → *Differentiate*) график экспериментальных точек (рис. 16).

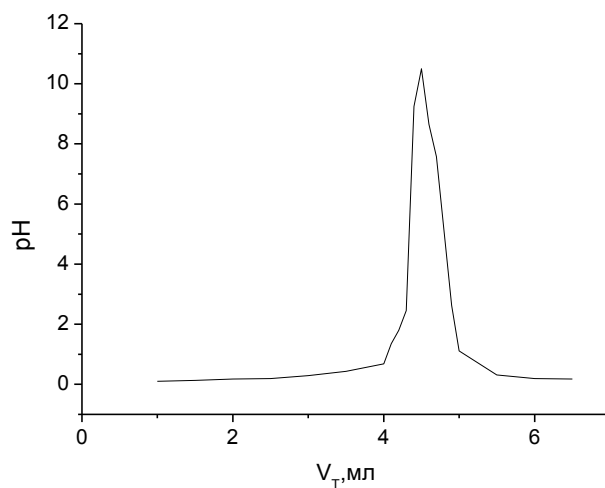


Рисунок 16 – Дифференциальная кривая титрования

4. Выделите данные листа **Derivative1- Derivative of Data1_B** и постройте новый график (рис. 17) дифференцирования экспериментальных данных (*Plot → Scatter*).

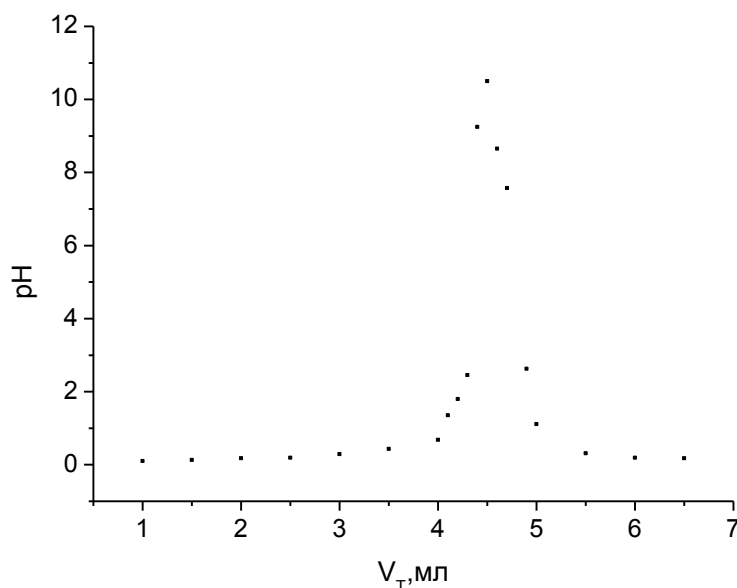


Рисунок 17 – Дифференциальная кривая титрования

5. Объедините два построенных графика, разместив их в разных слоях (*Layers*) одного графика. Для этого вернитесь кривой Graph1 (рис.1). В верхнем левом углу белого поля виден значок с цифрой 1, обозначающей номер единственного слоя графика. Добавьте второй слой (*Tools → Layer*). В появившемся окне **Layer** нажмите значок с выделенной правой вертикальной осью и закройте окно. На графике появится справа ось Y и слева вверху значок с цифрой 2, соответствующий новому слою.

6. Новый слой пока пуст, заполним его. Двойным щелчком на значке «2» откройте окно **Layer2**. В разделе *Available Data* выберите *derivative1data1b* - данные результатов дифференцирования, нажав ☐ >, скопируйте их в раздел *Layer Contents* и нажмите ОК. Результаты дифференцирования появятся на графике.

7. *Построение аппроксимационных кривых (рис.18)*. Размещение данных в разных слоях позволяет обрабатывать их по отдельности.

Выделите слой 1, щелкнув на значке «1» и постройте кривую (*Analysis → Fit Sigmodial*). Выделите слой данных дифференцирования «2» и постройте кривую Гаусса (*Analysis → Fit Gaussian*).

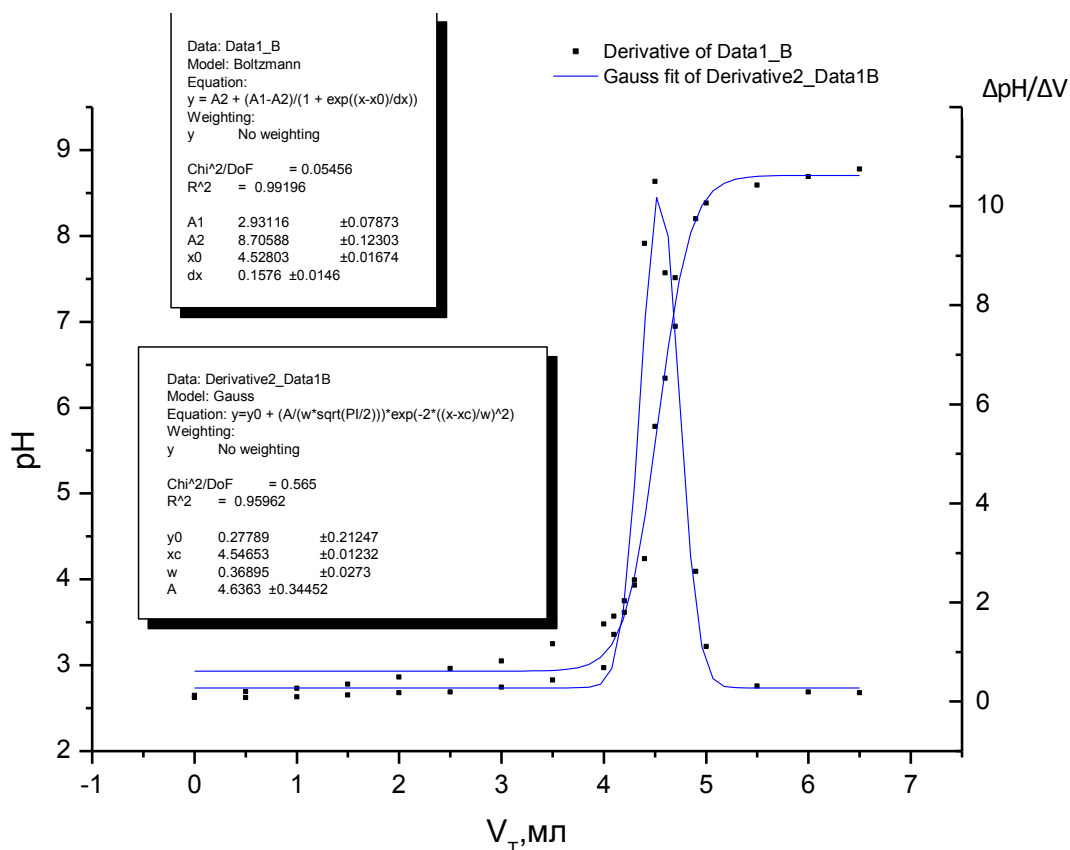


Рисунок 18 – Графическое представление результатов анализа

Задача

При оценке марганца в природной воде получено радиоактивационным методом (нг/мл): 0,35; 0,75; 0,65; 0,45; 0,55; 0,30; 0,80. Рассчитайте среднее и его доверительный интервал (при $p=0,95$). Оцените воспроизводимость определения.

Решение

1. Определение среднего значения

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$$

$$\bar{X} = \frac{0,35 + 0,75 + 0,65 + 0,45 + 0,55 + 0,30 + 0,80}{7} = 0,55$$

2. Определение дисперсии

Дисперсия S^2 характеризуется рассеиванием результатов анализа относительно среднего значения.

$$S^2 = \frac{\sum x_i - \bar{x}^2}{n-1}$$

Выборочное стандартное отклонение отдельного определения S рассчитывают по формуле:

$$S = \sqrt{\frac{\sum x_i - \bar{x}^2}{n-1}}$$

Относительное стандартное отклонение S_r

$$S_r = \frac{S}{\bar{x}}$$

$$S^2 = \frac{(0,35-0,55)^2 + (0,75-0,55)^2 + (0,65-0,55)^2 + (0,45-0,55)^2 + (0,30-0,55)^2 + (0,8-0,55)^2}{7-1} = 0,0375$$

$$S = \sqrt{0,0375} = 0,193649$$

$$S_r = \frac{0,193649}{0,55} = 0,352089 \approx 0,35$$

$$\bar{x} \pm \Delta\bar{x} = \bar{x} \pm \frac{t_{p,f} \cdot S}{\sqrt{n}}$$

$$\Delta\bar{x} = \frac{t_{p,f} \cdot S}{\sqrt{n}}$$

$$t_{p,f} = 2,45$$

$$\Delta\bar{x} = \frac{2,45 \cdot 0,1936}{\sqrt{7}} = \frac{0,4743}{2,65} = 0,18$$

$$\bar{x} \pm \Delta\bar{x} = 0,55 \pm 0,18$$

7. Представление результатов определений

При выполнении нескольких определений одного и того же элемента в одном и том же образце экспериментатор получает несколько отличающихся друг от друга результатов (разброс результатов). Разность между наибольшим и наименьшим результатами дает числовое представление о качестве выполненных определений. Любое из нескольких полученных результатов определений не показывает истинного содержания элемента, так как оно обязательно включает в себя какие-то погрешности измерений. Фактически экспериментатор определяет не истинное содержание, а те пределы, в которых оно должно находиться с данной степенью точности.

Таблица 1 – Пример представления результатов анализа

Параметр	Обозначение	Пример записи
Исходные данные	x_i	4,83; 4,85; 4,87; 4,90
Число измерений	n	4
Однородность выборки	Q	$Q < (Q_{\text{крит.}})=0,76$
Среднее значение	$\bar{x}$	4,86
Дисперсия (средняя квадратичная погрешность)	S^2	0,009
Стандартное отклонение выборочной совокупности	S	0,03
Относительное стандартное отклонение для среднего значения	$S_r = \frac{S}{\bar{x}}$	0,006
Полуширина доверительного интервала для среднего значения	$\Delta\bar{x}$	0,05 ($P=0,95$)
Границы доверительного интервала	$\bar{x} \pm \Delta\bar{x}$	4,86 $\pm$ 0,05
Относительная погрешность для среднего значения	ε	1,03%

1. Проверка наличия грубых погрешностей:

Прежде чем приступить к обработке данных с применением методов математической статистики, необходимо выявить грубые погрешности (промахи) и исключить их из числа выборки. Для выявления промахов в ряде параллельных определений при небольшом числе измерений используют Q-критерий.

$$Q = \frac{x_1 - x_2}{W}, \text{ где } x_1 - \text{сомнительный результат; } x_2 - \text{соседний с ним}$$

по величине результат; W - размах варьирования, который равен разности между максимальным и минимальным значениями x в ряду параллельных результатов.

$$W = X_{\max} - X_{\min}$$

Рассчитанную Q величину сравнивают $Q_{\text{табл.}}$. Если рассчитанная величина $Q > Q_{\text{табл.}}$, подозрительный результат необходимо откинуть.

2. Дисперсия S^2 характеризуется рассеиванием результатов анализа относительно среднего значения.

$$S^2 = \frac{\sum x_i - \bar{x}^2}{n-1}$$

3. Выборочное стандартное отклонение отдельного определения S рассчитывают по формуле:

$$S = \sqrt{\frac{\sum x_i - \bar{x}^2}{n-1}}$$

4. Для характеристики воспроизводимости определяют относительное стандартное отклонение S_r

$$S_r = \frac{S}{\bar{x}}$$

5. Доверительный интервал среднего значения

$$\Delta \bar{x} = \frac{t_{p,f} \cdot S}{\sqrt{n}}$$

6. Статистическая надежность P или Доверительный интервал выборочного среднего арифметического рассчитывают по формуле

$$\bar{x} \pm \Delta \bar{x} = \bar{x} \pm \frac{t_{p,f} \cdot S}{\sqrt{n}}$$

$\bar{x}$ – среднее арифметическое выборки

$\Delta \bar{x}$ – точность погрешности среднего арифметического выборки

S – стандартное отклонение выборки

$t_{p,f}$ – коэффициент распределения Стьюдента (табл. знач.)

n – число вариантов (определений).

7. Точность анализа (относительная погрешность для среднего значения) рассчитывается по формуле

$$\varepsilon = \frac{\Delta \bar{x}}{\bar{x}} = \frac{t_{p,f} \cdot S}{\sqrt{n}} / \bar{x} \cdot 100\%$$

Так как истинное значение определяемой величины практически определить невозможно, то находят границы, в которых оно должно заключаться при данных условиях. Чем уже эти границы, тем выше

надежность и достоверность анализа. Для нахождения погрешности среднего значения результата его стандартное отклонение умножают на коэффициент Стьюдента (из табл. 1), называемый коэффициентом нормированных отклонений (при малом числе наблюдений), который выбирают из заданной доверительной вероятности.

ЛИТЕРАТУРА

1. Брановицкая С.В. Вычислительная математика в химии и химической технологии / С.В. Брановицкая, Р.Б. Медведев, Ю.Я. Фиалков. – К.: Вища школа, 1986. – с. 216.
2. Основы аналитической химии. В 2 кн. Кн. 1. Общие вопросы. Методы разделения / Под ред. Ю. А. Золотова. – 3-е изд., М.: Высш. шк., 2004. – 361 с.
3. Сайт компании-разработчика программного пакета Origin [Электронный ресурс]. URL: <http://www.originlab.com/> (Дата обращения 22.02.2018).
4. Новый справочник химика и технолога. Аналитическая химия. Часть I - СПб: "Мир и Семья", 2002. – 964 с.
5. ГОСТ Р 8.736-2011 Государственная система обеспечения единства измерений (ГСИ). Измерения прямые многократные. Методы обработки результатов измерений. Основные положения.

ПРИЛОЖЕНИЕ

Таблица 1 - Коэффициенты Стьюдента для различных чисел степеней свободы f и значений доверительной вероятности P

f	$P = 0.90$	$P = 0.95$	$P = 0.99$
1	6.31	12.71	63.66
2	2.92	4.30	9.93
3	2.35	3.18	5.84
4	2.13	2.78	4.60
5	2.02	2.57	4.03
6	1.94	2.45	3.71
7	1.90	2.37	3.50
8	1.86	2.31	3.36
9	1.83	2.26	3.25
10	1.81	2.23	3.17
11	1.80	2.20	3.11
12	1.78	2.18	3.06
13	1.77	2.16	3.01
14	1.76	2.15	2.98
15	1.75	2.13	2.95
16	1.75	2.12	2.92
17	1.74	2.11	2.90
18	1.73	2.10	2.88
19	1.73	2.09	2.86
20	1.73	2.09	2.85
30	1.70	2.04	2.75
40	1.68	2.02	2.71
60	1.67	2.00	2.66
∞	1.65	1.96	2.58

Таблица 2 - Коэффициенты Фишера для доверительной вероятности $P = 0.95$ и различных чисел степеней свободы f_1 и f_2

f_2	f_1													
	1	2	3	4	5	6	7	8	9	10	12	20	50	∞
1	161	200	216	225	230	234	237	239	241	242	244	248	252	254
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.41	19.44	19.47	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.74	8.66	8.58	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.80	5.70	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.68	4.56	4.44	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.87	3.75	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.57	3.44	3.32	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.28	3.15	3.03	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.07	2.93	2.80	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.91	2.77	2.64	2.54
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.69	2.54	2.41	2.30
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.12	1.96	1.84
50	4.04	3.19	2.80	2.57	2.41	2.30	2.23	2.14	2.08	2.04	1.96	1.80	1.61	1.45
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.57	1.36	1.00

Таблица 3 - Q-коэффициенты для доверительной вероятности $P = 0.90$ и различных значений n

n	Q	n	Q
3	0.94	7	0.51
4	0.76	8	0.47
5	0.64	9	0.44
6	0.56	10	0.41

СОДЕРЖАНИЕ

Введение	3
1. Погрешности. Классификация погрешностей химического анализа	4
1.1. По способу вычисления	4
1.2. По характеру причин, их вызывающих	5
1.3. Значащие цифры	6
1.4. Правильность. Методы проверки правильности	12
1.5. Воспроизводимость. Методы проверки воспроизводимости	13
1.6. Точность анализа	16
1.7. Исключение промахов из выборки	16
1.8. Сравнение двух или более средних значений	17
2. Расчет уравнения линейного градуировочного графика, его метрологических характеристик и метрологических характеристик результата анализа	17
3. Проверка коэффициента a уравнения регрессии на статистическую значимость по критерию Стьюдента	21
4. Проверка коэффициента a уравнения регрессии на статистическую значимость по F -критерию	21
5. Границы доверительного интервала для рассчитанного значения зависимой переменной	23
6. Компьютерная обработка результатов измерений	28
6.1. Построение градуировочного графика в программе Microsoft Office Excel	28
6.2. Построение градуировочного графика в программе OriginPro	33
7. Представление результатов определений	40
8. Литература	43
9. Приложение	44

Учебное издание

**Ткаченко Элла Владимировна
Толстенко Юлия Викторовна**

**МЕТРОЛОГИЧЕСКИЕ ОСНОВЫ
ХИМИЧЕСКОГО АНАЛИЗА**

Учебно-методическое пособие

В авторской редакции

Изд. № 283/19. Объем 3 п.л.
РИИЦМ ФГАОУ ВО «Севастопольский государственный университет»