

Статистика в аналитической химии.
Современные методы оценки
неопределенности.
Представление результатов.
Валидация аналитических методик.

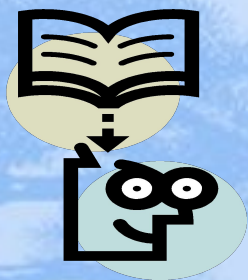
Плахотний Игорь Николаевич

*нач. отдела экспериментальных исследований, экологической безопасности
земель и качества продукции Днепропетровского проектно-технологического
центра*

+380964772507 , e-mail: igor_pin@mail.ru

URL: www.garryc2008.narod.ru

Статистика в аналитической химии



Определение понятия процесса измерения.

Химия, и в частности аналитическая химия, как и все точные науки опирается на измерение каких-либо величин. Процесс измерения, из-за своей распространенности представляется совершенно понятным действием. НО !

Всякое измерение невозможно без применения законов, относящихся к измерениям величин, и опирается на определенные теоретические предпосылки. Процесс получения количественной оценки измеряемых свойств объекта измерений опирается на несколько аксиом:

Аксиома 1. Измерение возможно при условии установления качественной определенности свойства, дающей возможность отличать его от других свойств, т.е. при условии выделения величины.

Аксиома 2. Измерение возможно при условии установления единицы, необходимой для измерения величины.

Аксиома 3. Измерение возможно при условии материализации (воспроизведения или хранения) единицы техническим средством.

Аксиома 4. Измерение возможно при условии сохранения неизменным размера единицы (в пределах установленной точности как минимум на срок, необходимый для измерений).

При несоблюдении хотя бы одного из этих условий, измерение не может быть выполнено. Аксиомы могут служить основой, на которую следует опираться, во-первых, при рассмотрении содержания понятия «измерение» и, во-вторых, при проведении четкой границы между измерением и другими видами количественного оценивания.

Статистика в аналитической химии

Использование статистики в химических, клинических и фармацевтических лабораториях является обычным делом.

Поскольку измеряемые данные получают в экспериментальных условиях, содержащих некоторую ошибку, статистические методы, хотя и не совершенны, являются наиболее эффективным путем осмысления данных.

Ситуация часто изображается, как :

$$T = t + e$$

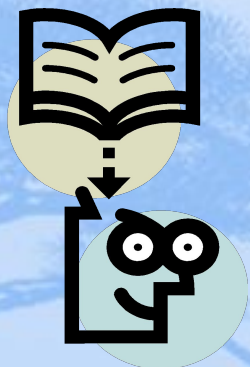
Здесь, истинное, но неизвестное значение, полученное в процессе измерения, T , состоит из результата измерения образца, t , и случайной ошибки или вариации, e , возникающей при выполнении этого измерения.

Статистическая ошибка является проявлением случайной изменчивости, свойственной любой системе измерений, и не является промахом. Например, инкубационная температура для бактерий в инкубаторе, имеет нормальное случайное колебание $\pm 1^\circ\text{C}$, которое рассматривается как статистическая ошибка. Статистическая ошибка – это не неправильное измерение или ошибочное измерение. Это, представление о неопределенности случайных колебаний.

Статистический анализ предоставляет возможность исследователю объяснить происхождение этой случайной ошибки.

Статистика в аналитической химии

Три типа погрешностей (ошибок)



В мире термин «погрешность» указывает на некую неудачу, промах. В метрологии погрешность определяется как «результат измерения минус истинное значение измеряемой величины» и свободно от такого негативного подтекста. Погрешность в химическом анализе - это конкретная величина, которая может быть известна, только если известно истинное значение.

Если мы проводим эксперимент, то мы почти всегда получаем результат с погрешностью. Почему мы не получаем истинное значение? Мы можем допустить погрешность при взвешивании, калибровке или даже при вычислениях. Повторный эксперимент может обнаружить эту погрешность. Первый тип погрешности, который можно назвать и ошибкой, называется грубая погрешность. А это действительно промах, и в случае такой погрешности ни какая книга по статистике или аналитической химии не в состоянии помочь её избежать (кроме простого совета - ☺ - «тщательнее работать надо»).

С учетом того, что мы сами можем выполнить наш анализ не совсем правильно, второй возможностью внесения погрешности является и тот факт, что методика сама по себе может иметь «червоточинку». В этом случае никакое количество повторов не может поправить ситуацию. Этот второй тип погрешностей называется систематической погрешностью и представляет собой постоянное отклонение от истинного значения. Систематическая погрешность может быть определена путем многократного измерения референсного материала (ОС). Разница между средним проведённых измерений и значением для референсного материала представляет собой систематическую погрешность. Всегда желательно знать источники систематической погрешности эксперимента и корректировать их при измерениях.

Статистика в аналитической химии

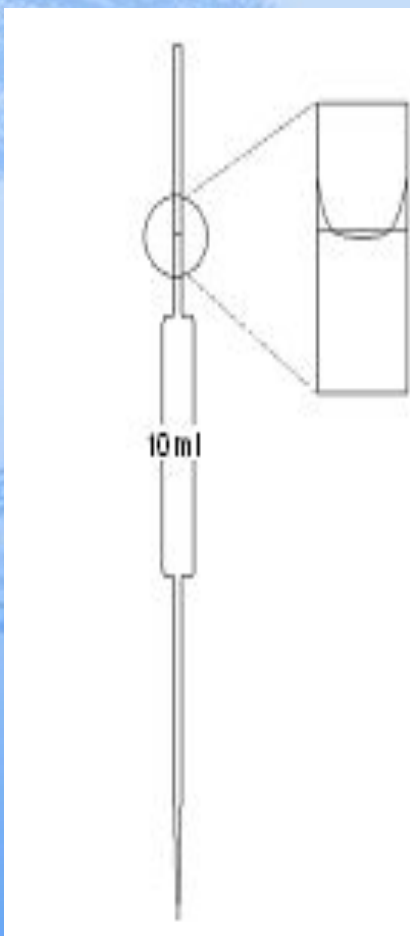
При оценивании величины систематической погрешности предполагается проводить эксперимент большое количество раз. Это необходимо еще и по причине вклада другого источника погрешности - случайной погрешности. Случайная погрешность это третий тип погрешности, который может быть ответственным за тот факт, что результат нашего эксперимента содержит погрешность даже после того как мы устранили грубые ошибки, повторяем эксперимент несколько раз, и все равно каждый раз имеем несколько иную величину. Подчас результат несколько больше ожидаемого, подчас несколько меньше.

Хорошей новостью в данном случае является то, что вычисляя среднее из большого числа результатов можно получить приемлемый ответ. Большие и меньшие значения взаимно нивелируются. Имеются мириады факторов, которые могут давать вклад в случайную погрешность: неспособность аналитика точно воспроизводить условия, флуктуации окружающей среды, округление при вычислениях, и, наконец, квантовая природа материи. Что не влияет на случайную погрешность так это изменения условий, вызывающих постоянный дрейф базовой линии прибора и старение хроматографической колонки (хотя если подумать, то и они могут опосредованно влиять на случайную погрешность).

Тип ошибок	Поведение
Грубые	Промахи
Систематические	Всегда одного и того значения и знака
Случайные	Нормально распределены с нулевым средним

Статистика в аналитической химии

Иллюстрация систематической погрешности



1. Производитель пипетки утверждает тот факт, что произведенная им пипетка, правильно заполненная до метки при $20\text{ }^{\circ}\text{C}$ гарантирует объем примерно между значениями 9.98 и 10.02 мл. Отметим, что погрешность этого рода - систематическая погрешность.

2. Когда вы заполняете пипетку, то заполняете ли Вы ее каждый раз точно до одного и того же уровня? Серия из 10 экспериментов заполнения пипетки дистиллированной водой и взвешивание дает нам диапазон от 9.95 до 10.04 мл. Вклад аналитика в погрешность по определению случайный (если аналитик имеет верный глаз и необходимые знания, в противном случае он может, например, совмещать с меткой объема не нижний, а например верхний край мениска).

3. Вы знаете, что в процессе Вашего экспериментирования, температура в Вашей лаборатории может колебаться между 19.2 и $23.1\text{ }^{\circ}\text{C}$ а, соответственно, что объем 10 мл воды увеличится на 0.0021 мл при повышении температуры воды на каждый градус Цельсия. Если эксперимент занимает много времени для того, чтобы дать температуре меняться случайным образом вокруг некоторого среднего, то эти изменения дадут вклад в результаты весового измерения (пункт 2). Кроме того, кроме того случая, когда точная температура не составляет точно $20\text{ }^{\circ}\text{C}$, это также приведет к систематической погрешности.

Статистика в аналитической химии

Случайные и систематические ошибки в аналитической химии вызываются множеством различных причин. Вот основные источники ошибок:

- Большинство исследуемых веществ надо рассматривать как неоднородные (негомогенные). Поэтому несколько небольших взятых из них частей — аналитических проб — могут не иметь одинакового состава. А значит, уже только по этой причине результаты анализа будут подвержены случайным колебаниям. Из-за неквалифицированного одностороннего отбора проб может отдаваться предпочтение отдельным компонентам, а в итоге — систематическое искажение состава пробы.
- Все необходимые для анализа измеряемые величины, такие, например, как масса осадка или светопоглощение окрашенного раствора, можно определить лишь с ограниченной точностью. Эта точность задается применяемым методом измерения, характером измеряемой величины, а часто и субъективными причинами. Если исключить показания неправильно отрегулированных измерительных приборов и иные подобные отказы, то ошибки чаще всего проявляются в форме случайных отклонений. Их надо минимизировать выбором подходящих условий измерений .
- В классических методах анализа часто исследуемые пробы подвергаются химическим реакциям, продукты которых характеризуются по виду, составу и массе. Обычно эти реакции рассматривают как равновесные, причем равновесие стремятся сдвинуть как можно дальше в сторону продуктов реакции. Несмотря на это, в ходе реакций возникают как случайные (например, колебания растворимости из-за различных концентраций растворяемых солей), так и систематические ошибки (например, из-за соосаждения). Задача аналитика состоит в том, чтобы подобрать для каждого конкретного случая наиболее подходящие реакции.

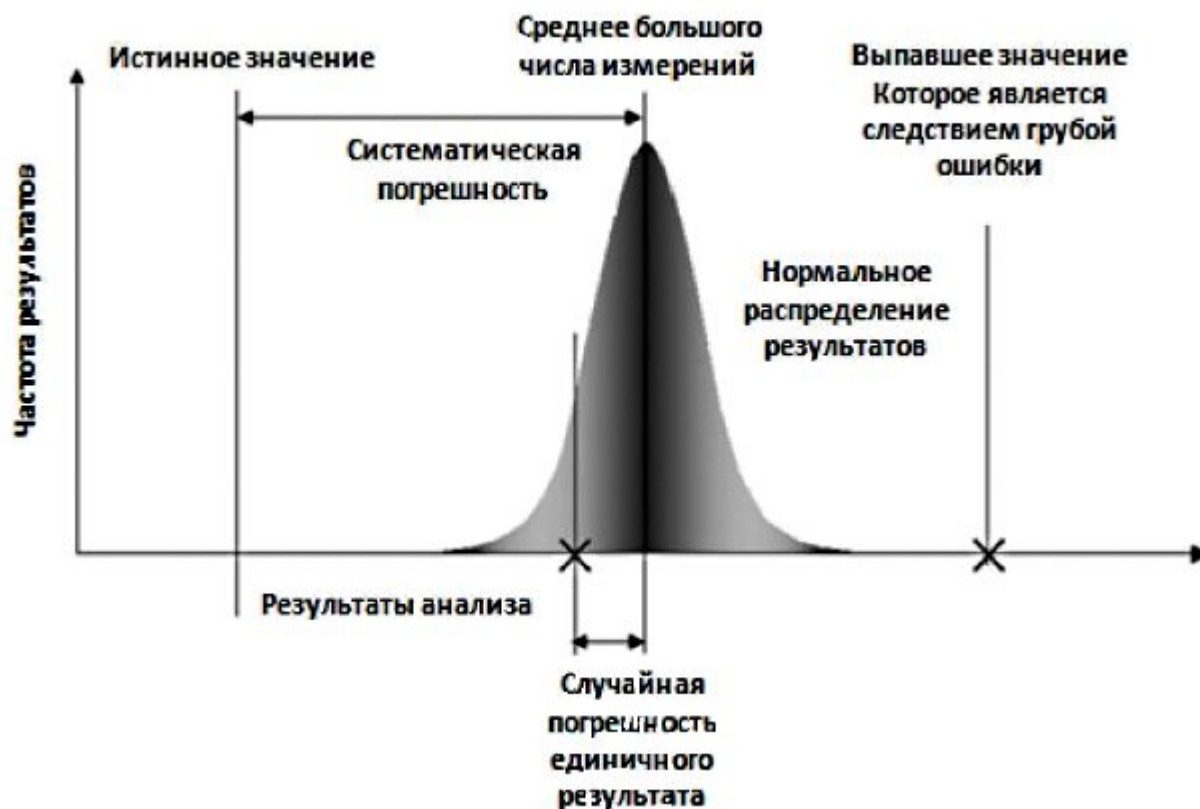
Статистика в аналитической химии

Если отбросить ошибку пробоотбора, как непосредственно не относящуюся к методу анализа, то общая ошибка складывается из ошибок измерений и ошибок, связанных с химическими реакциями. Как правило, ошибки измерений должны быть меньше, чем ошибки метода. В то время как ошибками измерений можно пренебречь, как это постоянно делается в физических исследованиях, для методических ошибок это не удастся или удастся лишь в исключительных случаях. Их описание, равно как и описание общей ошибки, возможно только с помощью методов математической статистики.

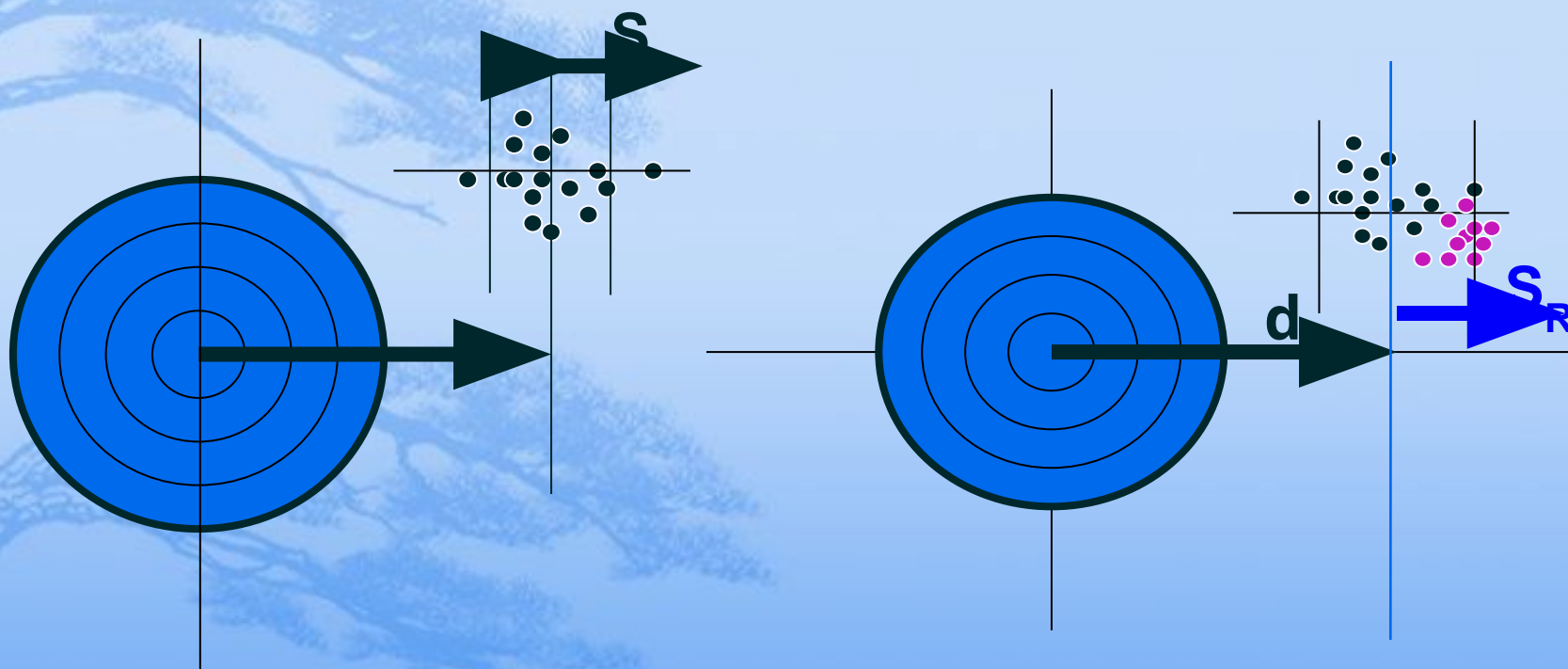
С понятиями систематической и случайной погрешности тесно связаны два важнейших метрологических понятия - **правильность и воспроизводимость**. **Правильностью** называется качество результатов измерения (или измерительной процедуры в целом), характеризующее величину систематической погрешности, **воспроизводимостью** - качество, характеризующее соответственно, значение случайной погрешности. Иными словами, правильность результатов - это их несмещенность, а воспроизводимость - их стабильность. Обобщающее понятие, характеризующее малость любой составляющей неопределенности - как систематической, так и случайной, - называется **точностью**. Мы называем результаты точными только в том случае, если для них мала как систематическая, так и случайная погрешность. Таким образом, правильность и воспроизводимость - это две составляющие точности, называемые поэтому **точностными характеристиками**.

В химической метрологии традиционно принято оценивать точностные характеристики по отдельности.

Статистика в аналитической химии

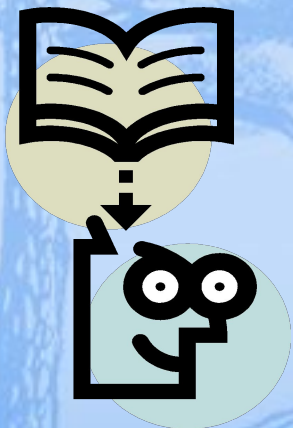


Статистика в аналитической химии



Статистика в аналитической химии

Основные статистические понятия в аналитическом измерении (генеральная совокупность и выборка, среднее, дисперсия, SD и RSD, степени свободы, доверительный интервал).



Генеральная совокупность и Выборка

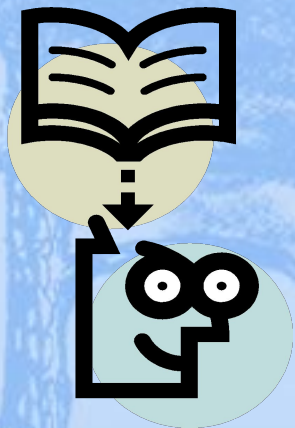
Статистики называют бесконечное число результатов, которые могли бы быть получены, в отношении которых формулируется исследовательская гипотеза, и которые описываются функцией распределения вероятности - генеральной совокупностью.

Выборка - это ограниченная по численности группа объектов, специально отбираемая из генеральной совокупности для изучения ее свойств.

Целью многих методов анализа данных является оценка μ и σ из всего нескольких повторных измерений, которые и называются выборкой (sample).

Здесь имеется небольшая проблема в обозначениях, которая ощущается только англоязычными химиками. Они привыкли называть то что анализируют «sample», но и статистика называет выборку тоже «sample», поэтому возникает путаница. International Union of Pure and Applied Chemistry (IUPAC) рекомендует называть «то что анализируют» так, сначала слово «test» затем слово-характеристика материала, например, «test solutions» или «test extract», или «test tablets», слово «sample» зарезервировано за статистикой.

Статистика в аналитической химии



Мера центральной тенденции любой выборки - это число, характеризующее выборку по уровню выраженности измеренного значения. Существуют три способа определения «центральной тенденции», каждому из которых соответствует своя мера: мода, медиана и , собственно, - выборочное среднее.

Наиболее просто получаемой мерой центральной тенденции является мода.

Мода - это такое значение из множества измерений, которое встречается наиболее часто. Моде, или модальному интервалу признака, соответствует наибольший подъем (вершина) графика распределения частот. Если график распределения частот имеет одну вершину, то такое распределение называется унимодальным.

$$M_0 = X_{M_0} + \frac{h \cdot (m_{M_0} - m_{M_0+1})}{(2 \cdot m_{M_0} - m_{M_0-1} - m_{M_0+1})}$$

X_{M_0} - Начало модального интервала

m_{M_0} -Частота модального интервала

m_{M_0+1} -Частота интервала, следующего за модальным

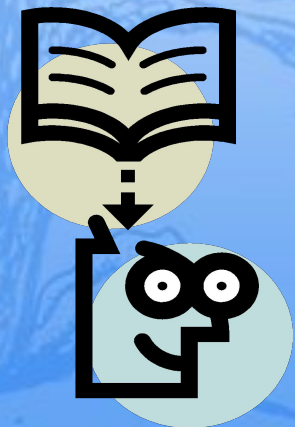
m_{M_0-1} -Частота интервала, предшествующего модальному

Статистика в аналитической химии

Медиана представляет собой срединное значение данных, расположенных в восходящем порядке. Если имеется нечетное число данных, то это единственная срединная величина. Если имеется четное число данных, то медиана представляет собой среднее из двух срединных значений. Она робастна (устойчива), так как не важно насколько отвратительными являются одно или более экстремальных значений, так как они являются значениями в самом начале или самом конце списка. Их величина незначительна.

Медиана - это такое значение, которое делит упорядоченное (ранжированное) множество данных пополам так, что одна половина всех значений оказывается меньше медианы, а другая - больше. Таким образом, первым шагом при определении медианы является упорядочивание (ранжирование) всех значений по возрастанию или убыванию. Далее медиана определяется следующим образом:

- если данные содержат нечетное число значений (8, 9, 10, 13, 15), то медиана есть центральное значение;
- если данные содержат четное число значений (5, 8, 9, 11), то медиана есть точка, лежащая посередине между двумя центральными значениями.



$$m_e = \begin{cases} X_{(m)}, n = 2m + 1, \\ \frac{1}{2} [X_{(m-1)} + X_{(m)}], n = 2m, \end{cases}$$

X_m – ранжированные данные

m - число измерений.

n - порядковый номер медианной варианты

Статистика в аналитической химии

Среднее

Выборочное среднее (арифметическое среднее) это результат суммирования всех результатов и деления на число данных (**n**):

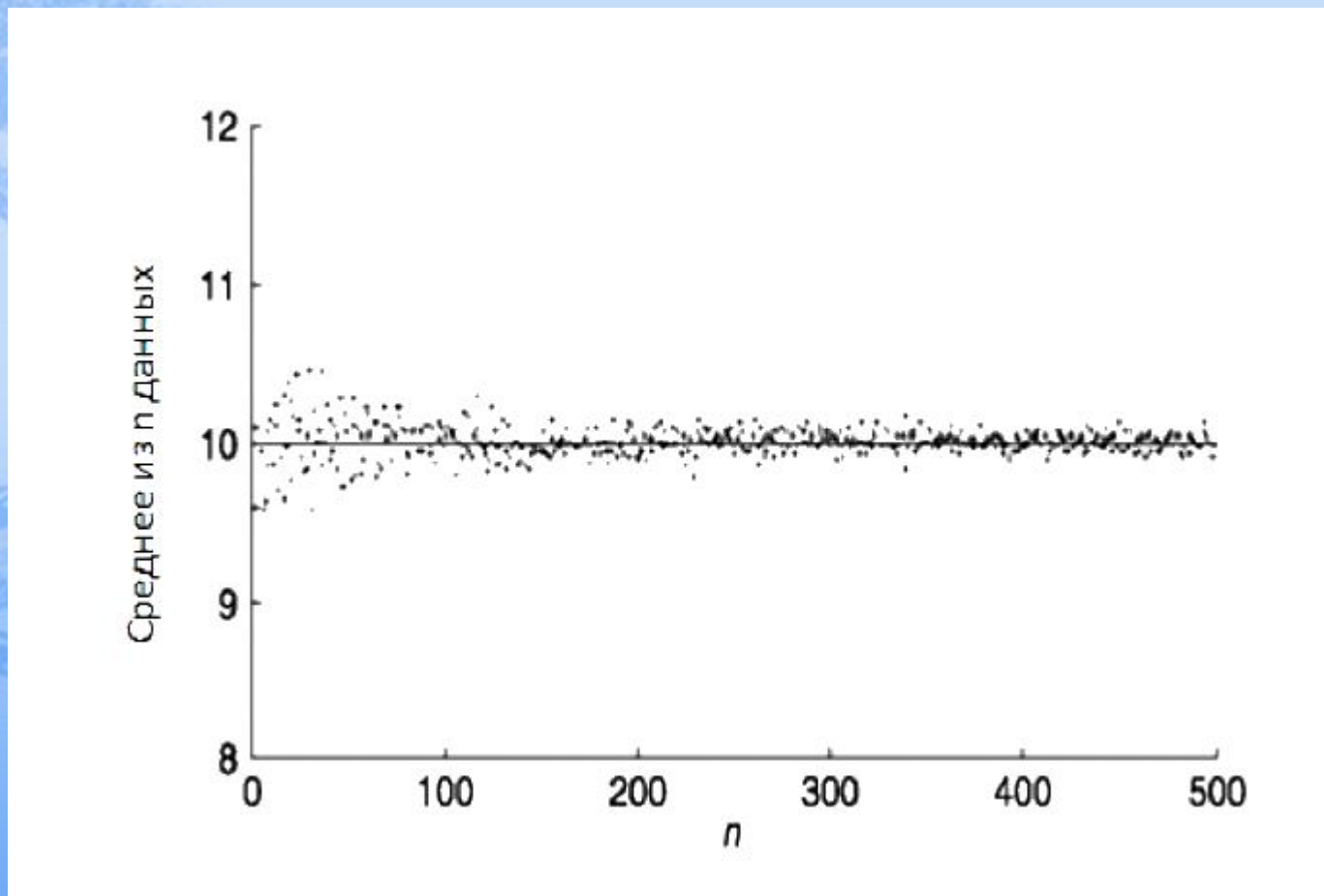
$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Для нормально распределенных данных, по мере роста числа данных выборочное среднее \bar{x} стремится к генеральному среднему μ . Генеральное среднее есть истинный результат при отсутствии систематической ошибки. Хотя единичный результат взятый из нормально распределенной генеральной совокупности вероятнее ближе к среднему.

Каждая мера центральной тенденции обладает характеристиками, которые делают ее ценной в определенных условиях.

Для номинальных данных, разумеется, единственной подходящей мерой центральной тенденции является мода, или модальная категория - та градация номинальной переменной, которая встречается наиболее часто. Для порядковых и метрических переменных, распределение которых унимодальное и симметричное, мода, медиана и среднее совпадают. Чем больше отклонение от симметричности, тем больше расхождение между значениями этих мер центральной тенденции. По этому расхождению можно судить о том, насколько симметрично или асимметрично распределение. Однако использование среднего ограничивается тем, что на его величину влияет каждое отдельное значение. Таким образом, среднее значение весьма чувствительно к «выбросам» - экстремально малым или большим значениям переменной.

Статистика в аналитической химии



Среднее из n данных выбранных случайно из нормально распределенной популяции с $\mu = 10$ а $\sigma = 1$ как функция от n

Статистика в аналитической химии

Реальные данные могут быть нормально распределены, но часто распределение содержит данные, которые могут быть серьезно сдвинуты.

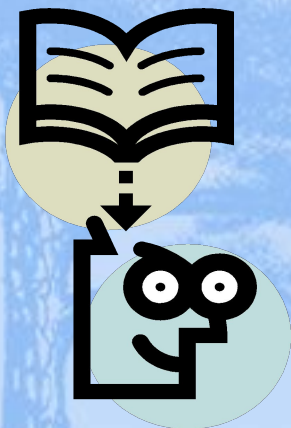
Если мы можем идентифицировать такие данные и удалить из последующего рассмотрения, то все хорошо. Иногда это возможно, иногда нет. Это может быть проблемой, потому что одно единственное негодное значение может серьезно расстроить вычисления среднего и стандартного отклонения. Оценки, которые могут сопротивляться определенному количеству плохих данных называются робастными и могут быть использованы в том случае, когда нет возможности обеспечить корректность характеристик обрабатываемых данных. Такими робастными оценками выступают - срединное значение упорядоченного набора данных (медиана) которая является робастной оценкой среднего и диапазон срединных 50 % данных (нормализованный интерквартильный диапазон) как робастная оценка стандартного отклонения. Робастные оценки имеют свою область применения, особенно тогда, когда необходимо сохранить все данные вместе, скажем в межлабораторных испытаниях, когда выпавший результат отдельной лаборатории нельзя просто так проигнорировать. Однако, робастные оценки не самые лучшие статистические характеристики и там, где можно, то следует использовать иные параметры распределения.

Интерквартильный диапазон (interquartile range (IQR)) это диапазон значений охватывающий 50% срединных данных. Три четверти IQR, называются нормализованным IQR и он служит для оценки стандартного отклонения.

Проблемой для IQR является то, что он не может быть вычислен для малых наборов данных, так как должно быть достаточно данных для определения квартилей (фрагментов упорядоченных данных, которые содержат одну четверть данных).

Интерквартильным размахом называется разность между третьей и первой квартилями, то есть $x_{0.75} - x_{0.25}$.

Статистика в аналитической химии



Стандартное отклонение и дисперсия

Разброс генеральной совокупности, демонстрируемый «тучностью» колоколообразной кривой измеряется посредством стандартного отклонения SD или его квадрата дисперсии. Они определяются для выборки из (**n**) данных как:

Дисперсия

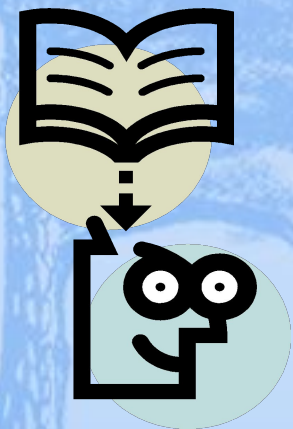
$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

СКО (SD (standard deviation))

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Стандартное отклонение **S**, определенное в уравнении известно как выборочное стандартное отклонение так как оно указывает на выборку из **n** значений и является оценкой популяционного стандартного отклонения **σ**.

Статистика в аналитической химии



Относительное стандартное отклонение

Относительное стандартное отклонение (RSD), известное также как коэффициент вариации (CV), это стандартное отклонение измерения выраженное в виде доли или процентной части от среднего:

$$RSD = \frac{S}{x} \cdot 100\%$$

Относительное стандартное отклонение аналитического результата представляется часто по причине того, что оно дает немедленное впечатление от точности измерения. Менее чем 1% обычно рассматривается как очень хорошее для рутинных измерений, для которых оно чаще находится в диапазоне 1 - 5%.

Стандартное отклонение среднего

Выборочное среднее тем ближе к генеральному, чем больше число данных в выборке. А в качестве выражения этой уверенности можно определить «стандартное отклонение среднего» (standard deviation of the mean). Для объяснения этого, предположим, что мы усредняем четыре точки, затем еще четыре и т.д. Среднее из четырех точек образует группу (на жаргоне другую «популяцию») со своим собственным средним и стандартным отклонением, которое может быть связано со средним и стандартным отклонением исходного набора данных из которых эта выборка была взята.

Статистика в аналитической химии

В математической статистике есть так называемая центральная предельная теорема, которая гласит о том, что среднее из средних является генеральным средним и стандартное отклонение популяции средних из n данных (σ_n) связано с генеральным стандартным отклонением (σ) формулой:

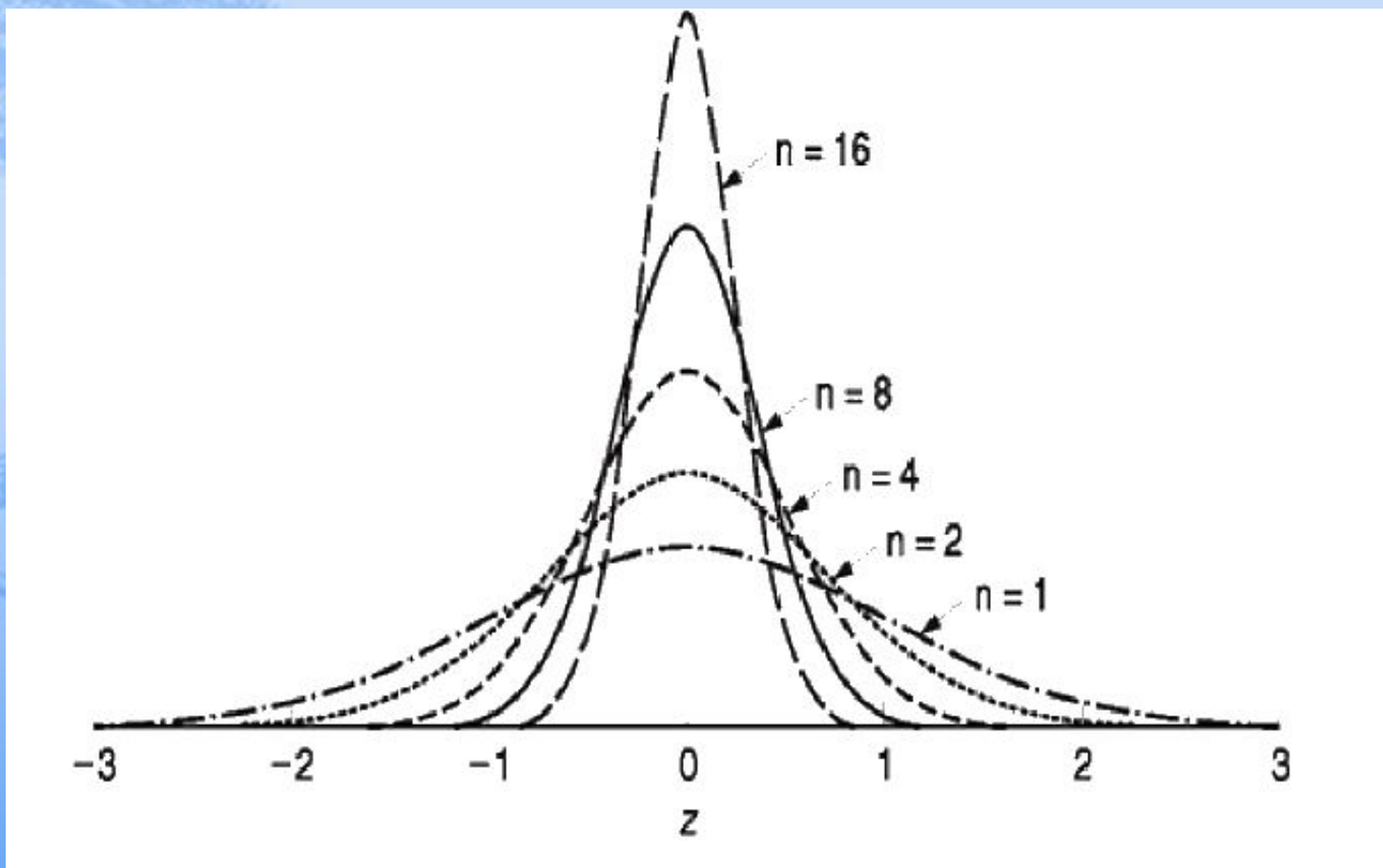
$$\sigma_n = \frac{\sigma}{\sqrt{n}}$$

Так называемое выборочное стандартное отклонение среднего определяется как :

$$S_n = \frac{S}{\sqrt{n}}$$

Стандартное отклонение среднего таким образом также становится все меньше по мере того как число данных в выборке возрастает, что отражает рост нашей уверенности в среднем значении .

Статистика в аналитической химии



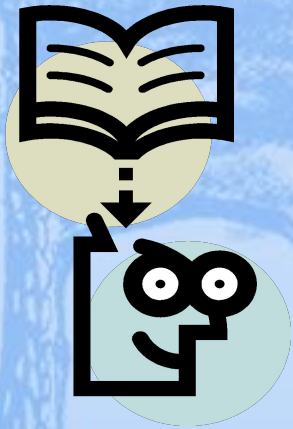
Статистика в аналитической химии

Несколько огорчает наличие корня квадратного от n в уравнениях. Этот корень означает, что, если хочется уменьшить стандартное отклонение в два раза, то надо увеличить число измерений в четыре раза. Число повторов зависит от того, где и как будет использоваться результат, т.е.

Центральная предельная теорема также дает аналитику еще один позитив. В большинстве случаев полученные химиком-аналитиком данные предположительно имеют нормальное распределение.

Но, к сожалению реальное распределение данных не так уж и нормально, но распределение средних по выборкам всегда стремится к нормальному распределению, даже если генеральная совокупность не является нормально распределенной. Таким образом нахождение средних по данным также помогает нам с анализом данных путем устранения озабоченности о том нормально или не нормально распределены данные в генеральной совокупности.

Статистика в аналитической химии



Число степеней свободы (*degrees of freedom (df)*)

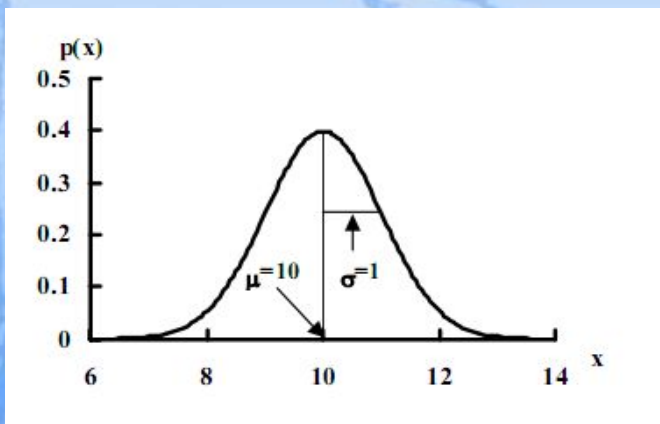
Число степеней свободы — одно из основных понятий в математической статистике. Это понятие трудно поддается строгому определению. В простых случаях число степеней свободы может быть интерпретировано как число переменных, которые могут быть произвольно присвоены при характеристике выборки. Другими словами, для выборки простейшей структуры (X_1, X_2, \dots, X_n) число степеней свободы равно разности между числом измерений и числом исследуемых параметров. В данном случае число измерений равно n , исследуемый параметр один (среднее значение), следовательно, число степеней свободы равно $df = n - 1$.

Доверительные интервалы и доверительные пределы

Стандартное отклонение среднего говорит нам однозначно о дисперсии (рассеивании) данных. Поэтому достаточно привести среднее значение и его стандартное отклонение (и число данных). Но более информативным является представление диапазона значений, которые охватывали бы некоторую часть данных (скажем 95 или 99%). Таким образом мы с некоторой долей скептицизма указываем конечный диапазон в котором с большой вероятностью находится истинное значение.

Статистика в аналитической химии

Строгий расчет границ доверительного интервала случайной величины возможен лишь в предположении, что эта величина подчиняется некоторому известному закону распределения. Закон распределения случайной величины - одно из фундаментальных понятий теории вероятностей. Он характеризует относительную долю (частоту, вероятность появления) тех или иных значений случайной величины при ее многократном воспроизведении. Математическим выражением закона распределения случайной величины служит ее функция распределения (функция плотности вероятности) $p(x)$.



Например, функция распределения, изображенная на сунке означает, что для соответствующей ей случайной величины x наиболее часто встречаются значения находятся вблизи $x = 10$, а большие и меньшие значения встречаются реже, чем дальше они отстоят от 10.

В качестве примера не случайно приведена колоколообразная, симметричная функция распределения. Именно такой ее вид наиболее характерен для результатов химического анализа. В большинстве случаев закон распределения результатов химического анализа можно удовлетворительно аппроксимировать так называемой функцией *нормального (или гауссова) распределения*:

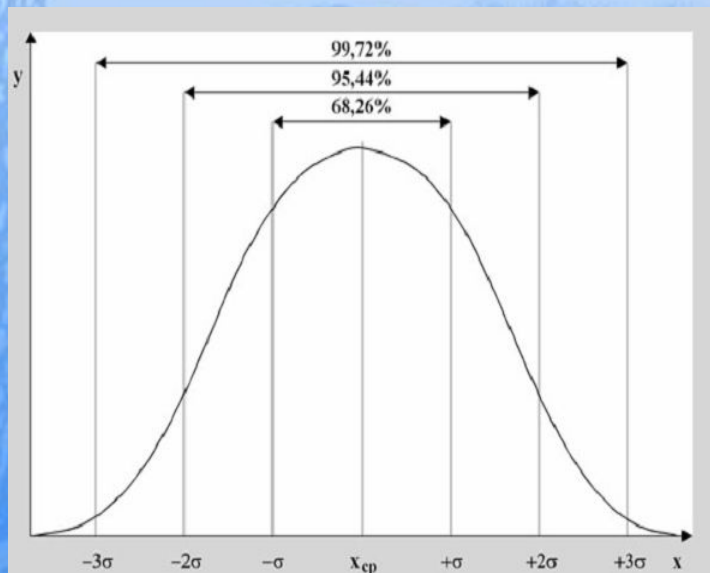
$$p(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}$$

Статистика в аналитической химии

Параметры этой функции μ и σ характеризуют: μ - положение максимума кривой, т.е. собственно значение результата анализа, а σ - ширину "колокола", т.е. воспроизводимость результатов. Можно сказать, что среднее является приближенным значением μ , а стандартное отклонение $s(x)$ - приближенным значением σ . Естественно, эти приближения тем точнее, чем больше объем экспериментальных данных, из которых они рассчитаны, т.е. чем больше число параллельных измерений n и, соответственно, число степеней свободы df . В предположении того факта, что, случайная величина x подчиняется нормальному закону распределения, то ее доверительный интервал рассчитывается как :

$$x \pm t(P, df) \cdot s(x)$$

Ширина доверительного интервала нормально распределенной случайной величины пропорциональна величине ее стандартного отклонения.



Площадь под кривой интерпретируется как вероятность или относительная частота.

Полезно знать, что если распределение является нормальным, то:

- 90% всех случаев располагается в диапазоне значений M (среднее) $\pm 1,64 \sigma$;
- 95% всех случаев располагается в диапазоне значений M (среднее) $\pm 1,96 \sigma$;
- 99% всех случаев располагается в диапазоне значений M (среднее) $\pm 2,58 \sigma$.

Статистика в аналитической химии

Численные значения коэффициентов пропорциональности t были впервые рассчитаны английским математиком В.Госсетом, подписывавшим свои труды псевдонимом Стьюдент, и потому называются коэффициентами Стьюдента. Они зависят от двух параметров: доверительной вероятности P и числа степеней свободы df , соответствующего стандартному отклонению $s(x)$. Причина зависимости t от P очевидна: чем выше доверительная вероятность, тем шире должен быть доверительный интервал с тем, чтобы можно было гарантировать попадание в него значения величины x . Поэтому с ростом P значения t возрастают. Зависимость t от df объясняется следующим образом. Поскольку $s(x)$ - величина случайная, то в силу случайных причин ее значение может оказаться заниженным. В этом случае и доверительный интервал окажется более узким, и попадание в него значения величины x уже не может быть гарантировано с заданной доверительной вероятностью. Чтобы "подстраховаться" от подобных неприятностей, следует расширить доверительный интервал, увеличить значение t - тем больше, чем менее надежно известно значение s , т.е. чем меньше число его степеней свободы. Поэтому с уменьшением df величины t возрастают.

Коэффициенты Стьюдента для различных значений P и df являются табличной величиной.

Если единичные значения x имеют нормальное распределение, то и среднее \bar{x} тоже имеет нормальное распределение. Поэтому формулу Стьюдента для расчета доверительного интервала можно записать и для среднего:

$$\bar{x} \pm t(P, df) \cdot s(\bar{x})$$

Статистика в аналитической химии

Величина $s(\bar{x})$ меньше, чем $s(x)$ (среднее точнее единичного). Для серии из n значений :

$$S(\bar{x}) = \frac{S(x)}{\sqrt{n}}$$

Поэтому доверительный интервал для величины, рассчитанной из серии n параллельных измерений, можно записать как :

$$\bar{x} \pm \frac{t(P, df) \cdot s(x)}{\sqrt{n}}$$

При расчете доверительного интервала встает вопрос о выборе доверительной вероятности P . При слишком малых значениях P выводы становятся недостаточно надежными. Слишком большие (близкие к 1) значения брать тоже нецелесообразно, так как в этом случае доверительные интервалы оказываются слишком широкими, малоинформативными. Для большинства химико-аналитических задач оптимальным значением P является 0.95. Именно эту величину доверительной вероятности (за исключением специально оговоренных случаев) мы и будем использовать в дальнейшем.

Статистика в аналитической химии

Зачем нужно проводить испытание гипотез?

*Неизменно помни, что природа - не бог,
человек - не машина, гипотеза - не факт.*

Дени Дидро

Статистика в аналитической химии

Одним из применений анализа полученных данных является возможность дать ответы на вопросы о качестве данных или о системе, которую эти данные описывают. В первую категорию попадают такие вопросы «распределены ли данные нормально?» и «имеются ли среди данных выпавшие значения?». Вопросы о системе могут быть такими: «Превышен ли уровень содержания примеси в лекарственном препарате?» или «дает ли новый метод такие же результаты как и традиционный метод?». Отвечая на эти вопросы мы определяем вероятность того, что представленные данные подтверждают правильность сформулированной гипотезы - следовательно «проверяем гипотезу».

Гипотеза, это утверждение, которое может быть истинным, а может быть ложным.

Обычно, гипотеза формулируется таким образом, чтобы имелась бы возможность вычислить вероятность (P) (или протестировать статистический показатель, вычисленный из этих данных) и затем принять решение о том, - должна ли гипотеза быть принята (высокий уровень P) или отвергнута (низкий уровень P). Частный случай испытания гипотез - тест на значимость. Для этого случая реально выдвигается гипотеза о том, что значимой разницы нет, а имеющаяся разница возникает от случайных эффектов: это называется нуль-гипотеза (H_0). Если вероятность, того что данные соответствуют выдвинутой нуль-гипотезе, попадает в заранее определенное значение (скажем 0.05 или 0.01), то гипотеза отвергается с этой вероятностью. Таким образом, $p < 0.05$ означает, что, если бы нуль-гипотеза была справедлива, то наблюдаемые значения статистического параметра, вычисленного из этих данных в некоем диапазоне, составляли бы менее чем 5% экспериментов. При использовании этого утверждения для проверки значимости какого-нибудь статистического показателя заранее должно быть принято решение о том уровне вероятности ниже которого нуль-гипотеза отвергается и делается заключение о значимости разности.

Статистика в аналитической химии

Статистическая значимость, или так называемый P - уровень значимости - это основной результат проверки статистической гипотезы. Говоря техническим языком, это вероятность получения данного результата выборочного исследования при условии, что на самом деле для генеральной совокупности верна нулевая статистическая гипотеза - то есть нет никакой связи. Иначе говоря, это вероятность того, что обнаруженная связь или зависимость носит случайный характер, а не является свойством совокупности. Именно статистическая значимость, P - уровень значимости является количественной оценкой надежности связи: чем меньше эта вероятность, тем надежнее связь.

Для решения вопроса о том, что такое - разумный уровень вероятности принятия или отвержения гипотезы, т.е., насколько значимо «значимое», следует рассмотреть два сценария, в которых может быть принято неправильное решение. Первый - мы отвергаем гипотезу, когда она на самом деле верна (это т.н. ошибка Типа I). В биологических науках ошибка Типа I часто называется ложно отрицательным результатом. Здесь заключение о том, что проверяемая разница находится вне допустимых пределов уровня, который можно ожидать на основе нормального распределения, соответствует нуль-гипотезе, когда на самом деле H_0 истинна. Второй сценарий противоположен этому, он состоит в том, что тест значимости приводит аналитика к ошибочному принятию H_0 , в то время как она ложна (ошибка Типа II). Ошибка Типа II знакома биологам (врачам, фармацевтам и прочим) как ложно положительный результат.

Статистика в аналитической химии

Lower-Tail
Односторонний тест

$$H_0 : A \geq B$$

$$H_A : A < B$$

H_0 должна быть
отвергнута
при $-\alpha$



**Критическое
значение**

Upper-Tail
Односторонний тест

$$H_0 : A \leq B$$

$$H_A : A > B$$

H_0 должна быть
принята



**Критическое
значение**

Two-Tail
Двусторонний тест

$$H_0 : A = B$$

$$H_A : A \neq B$$

H_0 должна быть
принята

H_0 должна быть
отвергнута
при $-\alpha/2$



**Критическое
значение**

Решение: Если расчетное значение
меньше табличного, то H_0 отвергают.

Решение: Если расчетное значение
больше табличного, то H_0 отвергают.

Решение: Если расчетное значение
больше или меньше табличного, то
 H_0 отвергают.

Статистика в аналитической химии

Проверка гипотез обычно проходит следующие этапы:

1. Определение используемой статистической модели. Здесь выдвигают некоторый набор предпосылок относительно закона распределения случайной величины и его параметров. Например, закон распределения нормальный, величины независимы и пр.
2. Формулируют H_0 и H_1 .
3. Выбирают критерий (критериальную статистику), который подходит к выдвинутой статистической модели.
4. Выбирают уровень значимости α в зависимости от требуемой надежности выводов.
5. Определяют критическую область для проверки H_0 . Если значение критерия попадает в эту область, то H_0 отклоняется. При условии, что H_0 верна, вероятность попадания в критическую область равна α . Вид этой области (односторонняя или двухсторонняя) зависит от принятой H_0 .
6. Рассчитывают значение выбранного статистического критерия для имеющихся данных.
7. Рассчитанное значение критерия сравнивают с критическим (иногда называемым табличным) и затем решают принять или отклонить H_0 .

Статистика в аналитической химии

Односторонние и двусторонние критерии проверки значимости

Если цель исследования состоит том, чтобы выявить различие параметров двух генеральных совокупностей, которые соответствуют различным ее естественным условиям, то часто неизвестно, какой из этих параметров будет больше, а какой меньше. Например, если интересуются вариативностью результатов в контрольной и экспериментальной выборках, то, как правило, нет уверенности в знаке различия дисперсий или стандартных отклонений результатов, по которым оценивается вариативность. В этом случае нулевая гипотеза состоит в том, что дисперсии равны между собой, а цель исследования — доказать обратное, т.е. наличие различия между дисперсиями. При этом допускается, что различие может быть любого знака. **Такие гипотезы называются двусторонними.**

Но иногда задача состоит в том, чтобы доказать увеличение или уменьшение параметра; например, средний результат в экспериментальной выборке выше, чем контрольной. При этом уже не допускается, что различие может быть другого знака. **Такие гипотезы называются односторонними.**

Критерии значимости, служащие для проверки двусторонних гипотез, называются двусторонними, а для односторонних — односторонними.

Статистика в аналитической химии

Статистические критерии - это инструмент, для того чтобы иметь возможность ответить на вопросы типа :

- Как точно выполнены анализы?
- Сколько анализов мне придется сделать, чтобы преодолеть проблемы неоднородности образцов или аналитические ошибки?
- Продукт соответствует спецификации?
- С какой уверенностью могу утверждать, что предельное содержание превышено?
- Какой метод точнее и на нем следует остановиться (критерии на основании которых был выбран тот или иной метод) ?

Возникает вопрос о том, какой из критериев следует выбирать в том или ином случае. Ответ на этот вопрос находится за пределами формальных статистических методов и полностью зависит от целей исследования. Ни в коем случае нельзя выбирать тот или иной критерий оценивания после проведения эксперимента, поскольку это может привести к неверным выводам.

Если до проведения эксперимента допускается, что различие сравниваемых параметров может быть как положительным, так и отрицательным, то следует использовать двусторонний критерий.

Статистика в аналитической химии

Если же есть дополнительная информация, например, из предшествующих экспериментов, на основании которой можно сделать предположение, что один из параметров больше или меньше другого, то используется односторонний критерий.

Когда имеются основания для применения одностороннего критерия, его следует предпочесть двустороннему, потому что односторонний критерий полнее использует информацию об изучаемом явлении и поэтому чаще дает правильные результаты.

Риск одностороннего критерия в том, что он может назвать значимой переменную, которая не является значимой на самом деле. Односторонний критерий – это шанс назвать вашу переменную значимой, когда двусторонний критерий не срабатывает.

Двусторонние тесты более строгие в отличие от односторонних. Коэффициент может быть незначим при двустороннем тесте и значим при одностороннем, поэтому использование односторонних тестов может оказаться полезным, так как хочется иметь значимые коэффициенты.

Если односторонний тест не позволил отвергнуть нулевую гипотезу, то есть значимость коэффициента обосновать не удалось, то более строгий двусторонний тест также не отвергнет нулевую гипотезу, и коэффициент является незначимым.

Статистика в аналитической химии

Выбор подходящего статистического метода для проверки гипотезы

Критерий t-Стьюдента для одной выборки

Данный метод позволяет проверить гипотезу о том, что среднее значение изучаемого признака отличается от некоторого известного значения. Исходное предположение – распределение признака в выборке приблизительно соответствует нормальному.

Критерий t-Стьюдента для независимых выборок

Данный метод сравнения позволяет проверить гипотезу о том, что средние значения двух генеральных совокупностей, из которых извлечены сравниваемые независимые выборки, отличаются друг от друга.

Исходные предположения :

- 1) одна выборка извлекается из одной генеральной совокупности, а другая выборка, независимая от первой, извлекается из другой генеральной совокупности;
- 2) распределение признака в обеих выборках приблизительно соответствует нормальному;
- 3) дисперсии признака в 2-х выборках примерно одинаковы (гомогенны).

Альтернатива методу – непараметрический U-критерий Манна-Уитни (если распределение признака хотя бы в одной выборке отличается от нормального или дисперсии статистически достоверно различаются).

Результатом данного анализа будет наличие или отсутствие достоверного различия между двумя выборками измерений, учитывая, конечно, уровень достоверности ($p < 0,05$).

Статистика в аналитической химии

Критерий t-Стьюдента для зависимых выборок

Этот метод позволяет проверить гипотезу о том, что средние значения двух генеральных совокупностей, из которых извлечены сравниваемые зависимые выборки, отличаются друг от друга. Зависимая выборка – когда определенный признак измерен на одной и той же выборке дважды, например, до и после воздействия, лечения и т.п.

Исходные предположения:

- 1) каждому представителю одной выборки поставлен в соответствие представитель другой выборки;
- 2) данные двух выборок положительно коррелируют;
- 3) распределение признака в обеих выборках приблизительно соответствует нормальному.

Альтернатива методу – непараметрический критерий T - Вилкоксона (если распределение признака хотя бы в одной выборке отличается от нормального и t-критерий Стьюдента для независимых выборок (если данные для двух выборок не коррелируют положительно)).

Сравнение дисперсий 2-х выборок по критерию Фишера.

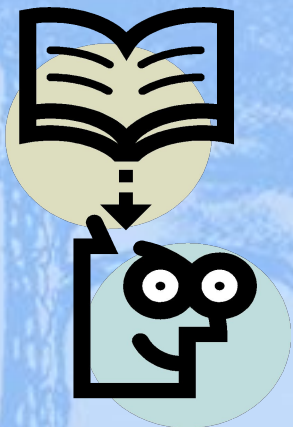
Данный метод позволяет проверить гипотезу о том, что дисперсии 2-х генеральных совокупностей, из которых извлечены сравниваемые выборки, отличаются друг от друга. Ограничения метода - распределения признака в обеих выборках не должны отличаться от нормального.

Альтернативой сравнения дисперсий является критерий Ливена, для которого нет необходимости в проверке на нормальность распределения.

Данный метод может применяться для проверки предположения о равенстве (гомогенности) дисперсий перед проверкой достоверности различия средних по критерию Стьюдента для независимых выборок разной численности.

Формулировка задачи в прикладной постановке	Формулировка задачи в статистической постановке	Дополнительные условия		Применяемый метод
Сравнение показателей контрольной и экспериментальной выборок	Проверка гипотезы о равенстве средних (центров распределения) в двух независимых выборках	Нормальный закон распределения	Дисперсии выборок равны	t-критерий (Стьюдента) при равных дисперсиях
			Дисперсии выборок не равны	t-критерий (Стьюдента) при неравных дисперсиях
		Без предположений о дисперсиях (но при одинаковом размере выборок)	t-критерий (Стьюдента) без предположений о дисперсиях	
		Дисперсии выборок равны	Манна - Уитни (U-критерий Вилкоксона - Манна - Уитни)	
		Без предположений о дисперсиях	Двухвыборочный Вилкоксона, медианный	
Сравнение показателей выборки до и после эксперимента	Проверка гипотезы о равенстве средних в двух зависимых выборках	Нормальный закон распределения		t-критерий (Стьюдента) для связанных выборок
		Закон распределения отличный от нормального или данные измеряются в дискретной шкале		Знаковый, одновыборочный критерий Вилкоксона
Можно ли считать, что среднее значение показателя равно некоторому номинальному значению?	Проверка гипотезы о равенстве среднего константе	Нормальный закон распределения		t-критерий (Стьюдента)
		Закон распределения отличный от нормального или данные измеряются в дискретной шкале		Гупта, знаковый
Сравнение рассеивания показателя в двух выборках	Проверка гипотезы о равенстве дисперсий (о принадлежности дисперсий к одной генеральной совокупности)	Нормальный закон распределения		F-критерий (Фишера)
		Закон распределения отличный от нормального или данные измеряются в дискретной шкале		Зигеля - Тьюки, Мозеса
Можно ли считать, что в нескольких выборках имеет место одно и то же значение показателя?	Проверка гипотезы о равенстве средних (о принадлежности средних к одной генеральной совокупности)	Нормальный закон распределения		Шеффе, Диксона, дисперсионный анализ, LSD
		Закон распределения отличный от нормального или данные измеряются в дискретной шкале		Краскела - Уоллиса, медианный, ранговых сумм Фридмана
Можно ли считать, что в нескольких выборках имеет место одно и то же значение рассеивания показателя?	Проверка гипотезы о равенстве дисперсий (о принадлежности дисперсий к одной генеральной совокупности)	Нормальный закон распределения		G-критерий (Кохрена) при равном размере выборок, Барглета
		Закон распределения отличный от нормального или данные измеряются в дискретной шкале		Фридмана

Статистика в аналитической химии



Что такое метод ANOVA ?

ANOVA (Analysis of Variance) – это дисперсионный анализ.

ANOVA это рабочая лошадка использования статистики для сравнения средних и определения эффектов влияния факторов на результаты измерения (т.е., чего угодно, что может изменяться или измеряться и при этом влиять на результаты эксперимента).

ANOVA, может разрешить вопрос о наличии значимого эффекта, вызываемого фактором для которого мы имеем какое-то количество наборов данных. ANOVA основывается на понимании двух вещей. Во-первых, того как дисперсии различных компонентов могут комбинироваться, давая общую дисперсию данных. Во-вторых, разница средних может привести к увеличению разброса комбинированных результатов, что может детектироваться в терминах возрастания дисперсии.

Слово «фактор» используется для указания на качество, которое исследуется. Примером может служить исследование влияния изменения полярности растворителя на результаты ВЭЖХ анализа. Фактором в данном случае является полярность растворителя, и мы можем сделать измерения, например, используя три растворителя с различной полярностью. В этом случае мы используем одно-факторный ANOVA, так как имеется только один исследуемый фактор - полярность растворителя.

То насколько изменяется измеряемая величина при изменении фактора имеет название «эффект» фактора. Часто в ANOVA мы всего лишь интересуемся проверкой того, имеется ли вообще какой-либо эффект.

Статистика в аналитической химии

Однофакторный ANOVA

В однофакторном ANOVA имеются уровни исследуемого фактора с повторными результатами для этих уровней. Например, мы можем оценить качество работы трех лабораторий, которые должны выполнить двукратный анализ идентичных образцов.

Таким образом, имеется один фактор - «лаборатория» для которого мы имеем три уровня - лаборатория А, лаборатория В и лаборатория С. Данные (результат каждого измерения) размещаются в матрицу с уровнями фактора в каждом столбце и повторами с каждой строке.

Общая планировка представления данных в однофакторном ANOVA

Повторы $i=(1\dots n_j)$	Уровни фактора $j=(1\dots k)$			
	1	2	...	k
1	$X_{1,1}$	$X_{1,2}$...	$X_{1,k}$
2	$X_{2,1}$	$X_{2,2}$...	$X_{2,k}$
...	$X_{i,j}$...
n_j	$X_{n_1,1}$	$X_{n_2,2}$...	$X_{n_j,k}$

Статистика в аналитической химии

Конкретная планировка представления данных ANOVA в межлабораторных сличительных испытаниях

Повторы	Лаборатория		
	А	В	С
1			
2			

Статистика в аналитической химии

Этапы в вычислении ANOVA таковы:

<p>1. Вычисляют среднее из всех данных: \bar{x} Где N - число результатов измерения. Это среднее называется общее среднее (grand mean). Вычитают \bar{x} из каждого результата измерения, так, что $x_{i,j}$ становится $(x_{i,j} - \bar{x})$. Эта операция известна как коррекция на среднее, а значения, как следствие, называются «значения, скорректированные на среднее».</p>	$\bar{x} = \frac{\sum_j \sum_i x_{i,j}}{N}$ $N = \sum_j n_j$
<p>2. Возводят в квадрат каждое из значений, скорректированных на среднее и суммируют их для получения полной суммы квадратов, известной как скорректированная сумма квадратов.</p>	$SS_T = \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{i,j} - \bar{x})^2$
<p>3. Для каждой колонки (т.е. для уровня фактора) усредняют значения скорректированные на среднее.</p>	$= \sum_{i=1}^{n_j} \frac{(x_{i,j} - \bar{x})}{n_j}$
<p>4. Возводят это полученное среднее в квадрат.</p>	$= \left(\sum_{i=1}^{n_j} \frac{(x_{i,j} - \bar{x})}{n_j} \right)^2$

Статистика в аналитической химии

5. Умножают на число строк (n_j) для данной колонки	$= n_j \left(\sum_{i=1}^{n_j} \frac{(x_{i,j} - \bar{x})}{n_j} \right)^2$
6. Суммируют по колонкам (sum j=1 to k) для получения суммы квадратов вследствие изучаемого фактора: SS_c известно так же под названием сумма квадратов между условиями испытаний, или межколоночная сумма квадратов. SS_c связана с межфакторной дисперсией.	$SS_c = \sum_{j=1}^k n_j \left(\sum_{i=1}^{n_j} \frac{(x_{i,j} - \bar{x})}{n_j} \right)^2$
7. Вычисляют остаточную сумму квадратов. SS_r также называется «внутрифакторной суммой квадратов».	$SS_r = SS_T - SS_c$

Статистика в аналитической химии

Таким образом, результат расчетов представляют в виде следующей таблицы

Источник	Сумма квадратов	Степени свободы	Средний квадрат	F
Между факторами	SS_c	$k-1$	$\overline{SS_c} = \frac{SS_c}{k-1}$	$F = \frac{\overline{SS_c}}{\overline{SS_r}}$
Внутри фактора	SS_r	$N-k$	$\overline{SS_r} = \frac{SS_r}{N-k}$	
ИТОГ	SS_T	$N-1$		

где N- общее число данных, k – число уровней фактора

Статистика в аналитической химии

Остаточный средний квадрат SS_r является оценкой средней дисперсии результатов в пределах каждого уровня фактора. В случае повторяющихся аналитических результатов SS_r представляет собой дисперсию повторяемости.

ANOVA, таким образом, представляет собой полезный способ оценки точности измерения. Обычно нас интересуют различия среди наших факторов, например, «Какая лаборатория лучше сработала на межлабораторных сличительных испытаниях - А или лаборатория В?»

К сожалению SS_c сам по себе не сообщает нам ответ, так как этот критерий включает дисперсию измерения. Каждое измерение имеет свою неопределенность. Проверка на значимость различий требует от нас сравнения средних квадратов между факторами и в пределах факторов. Если нет эффекта то SS_c и SS_r будут одинаковыми. Таким образом если мы хотим выяснить уровень значимости между уровнями фактора и повторяемостью, т.е. « SS_c значимо больше чем SS_r ?», то необходимо определить связанную вероятность для (F- критерия) с соответствующими степенями свободы ($df1$, $df2$), где для ANOVA $df1$ - межфакторная степень свободы ($k-1$) и $df2$ это внутрифакторная степень свободы ($N - k$).

Хотя ANOVA и может определить значимость эффекта фактора, но если имеется более чем два уровня фактора, то ANOVA не может определить какой из уровней значимо дает вклад в разницу (через которую определяется значимость). Например, предположим, что мы используем ANOVA для решения вопроса о том, имеется или нет значимая разница среди дозируемых объемов для серии из 5 или 10 мл пипеток, путем взвешивания десяти выпущенных объемов от каждой пипетки.

Статистика в аналитической химии

Матрица данных в этом случае будет иметь пять колонок (каждая пипетка) и десять строк (повторы), т.е. всего 50 данных. Если ANOVA приводит к заключению, что имеется существенная разница, то неизвестно это одна пипетка отличается от остальных четырех, или более одной? Разумная идея выяснения, какая же пипетка или пипетки отличаются от остальных, может быть реализована т.н. методом наименьшей значимой разницы (least significant difference (LSD)). ANOVA дала внутрифакторное стандартное отклонение ($\sqrt{SS_r}$), которое в этом случае является повторяемостью дозирования и взвешивания объема (S_r). Разница между средними любых двух колонок данных, как предполагается, имеет стандартное отклонение $\sqrt{2}S_r/\sqrt{n}$ где n представляет собой число повторов. 95% доверительный интервал разности между любыми двумя средними составляет $\sqrt{2}t_{0.05, 45}S_r/\sqrt{10}$, где значение t получено для 45 степеней свободы ($N = 50 - 5 = 45$). Число степеней свободы 45 так как имеется 50 данных и каждый уровень фактора (которых всего 5) отнимает одну степень свободы. Значение $\sqrt{2}t_{0.05, 45}S_r/\sqrt{10}$ представляет собой LSD т.е. максимальная разность между средними, которая будет восприниматься как несущественная. Если средние по всем колонкам упорядочить в возрастающем порядке величины, то любая разность между последовательными средними большая чем LSD свидетельствует о наличии значимой разницы.

Статистика в аналитической химии

Некоторые примеры применения статистических вычислений к аналитической практике (Excell)



1. Проверка на нормальность распределения данных

Многие статистические параметры, используемые химиками – аналитиками, основаны на допущении, что данные распределены нормально.

Иногда, в случае стандартного распределения средних, данные имеют тенденцию к нормальному распределению согласно теории, но чаще всего мы просто скрещиваем пальцы и используем нормальное распределение. Хотя и нет метода, который бы по трем-четырем значениям смог дать вразумительное свидетельство о распределении, но все же имеются тесты по подтверждению нормального распределения для наборов данных для минимум 10 значений.

Критерий согласия Пирсона

Статистикой критерия Пирсона служит величина

$$\chi^2 = \sum_{j=1}^e \frac{(n_j - np_j)^2}{np_j} \quad \chi^2 \leq \chi^2_{\alpha}$$

где p_j - вероятность попадания изучаемой случайной величины в j -и интервал, вычисляемая в соответствии с гипотетическим законом распределением $F(x)$.

Статистика в аналитической химии

Критерий согласия Пирсона (χ^2) применяют для проверки гипотезы о соответствии эмпирического распределения предполагаемому теоретическому распределению $F(x)$ при объеме выборки ($n \geq 100$). Критерий применим для любых видов функции $F(x)$, даже при неизвестных значениях их параметров. В этом заключается его универсальность.

Использование критерия χ^2 предусматривает разбиение размаха варьирования выборки на интервалы и определения числа наблюдений (частоты) n_j для каждого из e интервалов.

Однако содержание какого-либо химического вещества или элемента в препарате не является дискретной величиной. Фактически оно (содержание) величина непрерывная. Исходя из этих соображений и определяют фактические границы интервалов $X_{i \min} \div X_{i \max}$. Следующий этап расчетов заключается в определении вероятности попадания нормальной случайной величины X_i с предполагаемыми средним значением X_{cp} и среднеквадратичным отклонением SD в заданные интервалы. Иногда для этого используется функция Лапласа для нормализованной случайной величины. Однако, пользуясь современными математическими пакетами, этого можно не делать. Так, в Excel существует функция **НОРМРАСП(x; x_{cp} ; σ ; Интегральная)**. Функция **НОРМРАСП(x; x_{cp} ; σ ; 1)** возвращает вероятность того, что нормально распределенная случайная величина при среднем x_{cp} и среднеквадратичном σ окажется не больше значения X_i . Используя ее рассчитывают значение интегральной функции на границах интервалов $F(X_{i \min}) \div F(X_{i \max})$.

Затем необходимо подсчитать общую сумму анализируемых наблюдений $n = \sum n_j$, ну и общую сумму наблюдаемых вероятностей. При неукоснительном использовании критерия согласия Пирсона общая сумма наблюдаемых вероятностей должна быть равна 1. В нашем случае этого не происходит. Это обусловлено применяемым для проверки на нормальность методом построения гистограмм. В связи с этим рассчитывают еще одну величину - *предполагаемое общее количество испытаний* n' . Эту величину определяют по формуле :

Статистика в аналитической химии

$$n^* = \frac{n}{\sum p_i}$$

Далее определяют вспомогательные данные для расчета наблюдаемого критерия согласия Пирсона χ^2 .

Теперь наблюдаемое значение критерия χ^2 необходимо сравнить в критическим значением. Уровень значимости, обычный для аналитических исследований принимают равным $\alpha = 0.05$.

Количество степеней свободы равно количеству наблюдаемых интервалов (в примере = 9 уменьшенному на количество определенных параметров (2, так как определили для выборки среднее и среднеквадратичное отклонение) и на 1: $k = 9 - 2 - 1 = 6$.

Для уровня значимости $\alpha = 0.05$ и числа степеней свободы $k = 6$ определяют критическое значение χ^2 - статистики используют либо табличные данные, либо функцию Excel ХИ2ОБР(α ; k). Условие $\chi^2_{\text{набл.}} \leq \chi^2_{\text{крит.}}$ выполняется, следовательно гипотезу о нормальности закона распределения для исследуемой выборки отвергать нет основания.

Статистика в аналитической химии

$\text{=КОРЕНЬ}((N17-M17^2/B17)/(B17-1))$
 =M17/B17
 =(A7+A8)/2
 $\text{=НОРМРАСП}(D7;\$B\$2;\$B\$3;1)$
 =A7-(A8-A7)/2
 =F7-E7
 =B7-H7

Содержание углерода, %, X_i	Количество наблюдений, n_i	Границы интервалов X_i		Интегральная функция распределения на границах интервалов $F(x)$		Вероятность попадания в интервал, p_i	$n \cdot p_i$	$n_i - n \cdot p_i$	$(n_i - n \cdot p_i)^2 / n \cdot p_i$	$n_i \cdot X_i$	$n_i \cdot X_i^2$			
		$X_{i \min}$	$X_{i \max}$	$F(X_{i \min})$	$F(X_{i \max})$									
0,14	8	0,135	0,145	0,007	0,026	0,020	5,8	2,2	0,804	1,12	0,1568			
0,15	18	0,145	0,155	0,026	0,081	0,055	16,1	1,9	0,214	2,7	0,405			
0,16	34	0,155	0,165	0,081	0,195	0,114	33,6	0,4	0,004	5,44	0,8704			
0,17	43	0,165	0,175	0,195	0,375	0,179	52,8	-9,8	1,819	7,31	1,2427			
0,18	69	0,175	0,185	0,375	0,587	0,212	62,5	6,5	0,686	12,42	2,2356			
0,19	51	0,185	0,195	0,587	0,776	0,189	55,7	-4,7	0,390	9,69	1,8411			
0,20	36	0,195	0,205	0,776	0,903	0,127	37,4	-1,4	0,050	7,2	1,44			
0,21	23	0,205	0,215	0,903	0,967	0,064	18,9	4,1	0,887	4,83	1,0143			
0,22	8	0,215	0,225	0,967	0,991	0,024	7,2	0,8	0,088	1,76	0,3872			
n=	290					$\sum p_i$=	0,984			χ^2=	4,943	\sum=	52,47	9,5931
n' =	295									$\chi^2_{\text{табл}}$=	12,592			

=B17/G17
 $\text{=СУММ}(J7:J15)$
 $\text{=ХИ2ОБР}(0,05;СЧЁТ(B7:B15)-2-1)$

Статистика в аналитической химии

Критерий Шапиро-Уилка.

Этот критерий является одним из наиболее мощных в большинстве случаев. Статистика W вычисляется по формуле :

$$W = b^2 / S^2$$

где, $S^2 = \sum (x_i - \mu)^2$, $b = \sum a_{n-i+1}(x_{n-i+1} - x_i)$, μ - среднее выборки, a_{n-i+1} некоторые константы

Для удобства расчета этого критерия была выведена полезная аппроксимация, позволяющая применить критерий Шапиро-Уилка без помощи таблиц. Для $\alpha = 0,05$ предлагается статистика

$$W_1 = \left(1 - n^{\frac{0,6695}{0,6518}} \right) \cdot \frac{B}{S^2}$$

$$B = \left\{ \sum_{j=1}^m a_j (x_{n-j} - x_j) \right\}^2; m = \left[\frac{n}{2} \right]; a_0 = \frac{0,899}{(n-2,4)^{0,4162}} - 0,02; a_j = a_0 \left[z + \frac{1483}{(3-z)^{10,845}} + \frac{71,6 \times 10^{-10}}{(1,1-z)^{8,26}} \right]; z = \frac{n-2j+1}{n-0,5}$$

Если $W_1 < 1$ то нулевая гипотеза нормальности распределения случайных величин отклоняется. В случае вычисления критерия согласно ГОСТ ИСО 5479-2002 его величину сравнивают с табличной и если величина рассчитанного критерия больше табличного, то гипотеза о нормальном распределении данных принимается с уровнем значимости 0,05.

Статистика в аналитической химии

Этот критерий применяется, когда выборка содержит малое количество наблюдений ($n \leq 30$).

Рассмотрим алгоритм применения этого критерия:

1. Проранжировать данные расчётной таблицы в неубывающем порядке.
2. Получить разности между крайними значениями.

Например, из самого большого по значению наблюдения вычитают самое наименьшее, затем из второго по величине – второе по наименьшему значению, и т. д. Таким образом, $\Delta_j = X_{n-j} - X_j$

3. Полученные разности Δ_j умножить на табличные коэффициенты, находимые в зависимости от числа наблюдений и порядкового номера разности.
4. Найти b - сумму умножений, полученную в пункте 3.
5. Найти СКО исходных данных.
6. Рассчитать величину критерия W по формуле: $W = b^2 / SD^2$

Если $W_{\text{расчит.}} > W_{\text{крит.}}$, то принимается нулевая гипотеза о нормальности распределения.

Статистика в аналитической химии

	A	B	C	D	E	F	G	H
1	Расчет критерия Шапиро-Уилка для проверки гипотезы о нормальном распределении с использованием аппроксимационной формулы							
2								
3								
4	n	Содержание углерода, %, X_i		j	$X_{n-j} \cdot X_j$	z_j	a_j	$a_j \cdot (X_{n-j} - X_j)$
5	1	0,14		1	0,08	0,941176	0,607631051	0,048610484
6	2	0,15		2	0,06	0,705882	0,346212521	0,020772751
7	3	0,16		3	0,04	0,470588	0,208099601	0,008323984
8	4	0,17		4	0,02	0,235294	0,101122659	0,002022453
9	5	0,18						
10	6	0,19						
11	7	0,20						
12	8	0,21						
13	9	0,22						
14								
15								
16	n =	9						
17	m =	4						
18	$a_0 =$	0,389888285						
19	$X_{sr} =$	0,18						
20	SD =	0,027386128						
							$\Sigma =$	0,079729673
							B =	0,006356821
							Аппроксимированный Критерий Шапиро - Уилка	
							$W_1 =$	7,1207
							$W_1 > 1$	

=B14-B6

=(\$B\$16-2*D6+1)/(\$B\$16-0,5)

$=\$B\$18*(F6+(1483/(3-F6)^{10,845}) + ((71,6*10^{(-10))}/(1,1-F6)^{8,26}))$

=G6*E6

=СУММ(H6:H9)

=H11^2

=СЧЁТ(B6:B14)

=ОКРВНИЗ(B16/2;1)

=СРЗНАЧ(B6:B14)

=СТАНДОТКЛОН(B6:B14)

$=(0,899/(B16-2,4)^{0,4162})-0,02$

$=(1-0,6695/(B16^{0,6518}))*H12/(B20^2)$

Статистика в аналитической химии

Графическая ранжит- процедура

Имеется и полезная графическая процедура для испытания нормальности набора данных, которая может быть реализована в Excel. Называется она Ранжит-методом (Rankit method).

Процедура, построена для n числа данных и заключается в следующих операциях :

1. Необходимо рассортировать данные в порядке возрастания.
2. Вычислить кумулятивную частоту каждого значения, т.е., сколько данных имеют равное или меньшее значение.
3. Вычислить нормализованную кумулятивную частоту = кумулятивная частота/ $(n + 1)$.
4. Вычислить значение нормальной функции плотности вероятности, связанной с нормализованной кумулятивной частотой для каждого значения ($= z$).
5. Построить график зависимости z от исходных данных.

Статистика в аналитической химии

Если данные распределены нормально, то график должен быть линейен. Очевидно, что выпавшие значения будут видны как точки на экстремумах оси x , т.е. при значениях много больше, чем можно предположить.

Данные от 25 участников ППТ по титрованию

Данные	Кумулятивная частота	Нормализованные значения	z
0.092	1	0.038462	-1.768825039
0.0936	2	0.076923	-1.426076872
0.1134	3	0.115385	-1.198379702
0.1138	4	0.153846	-1.020076233
0.1139	5	0.192308	-0.869423773
0.11415	6	0.230769	-0.736315917
0.1142	7	0.269231	-0.615141105
0.1143	9	0.346154	-0.395725296
0.1143	9	0.346154	-0.395725296
0.1144	11	0.423077	-0.194028142
0.1144	11	0.423077	-0.194028142
0.1145	12	0.461538	-0.096558615
0.1146	13	0.500000	-1.39214E-16
0.1148	14	0.538462	0.096558615
0.115	16	0.615385	0.293381232
0.115	16	0.615385	0.293381232
0.1152	17	0.653846	0.395725296
0.1153	18	0.692308	0.502402223
0.1155	19	0.730769	0.615141105
0.1158	20	0.769231	0.736315917
0.1177	21	0.807692	0.869423773
0.1219	22	0.846154	1.020076233
0.1222	23	0.884615	1.198379702
0.1556	24	0.923077	1.426076872
0.9083	25	0.961538	1.768825039

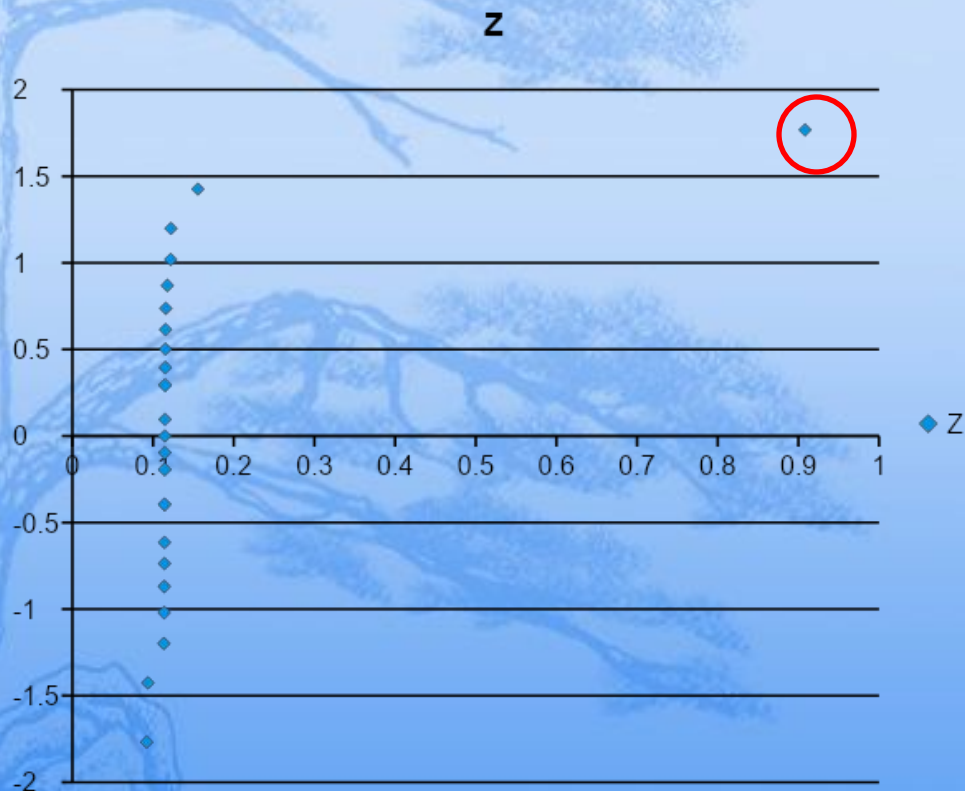
=СЧЁТ(\$A\$2:\$A\$26)+1-РАНГ(A2;\$A\$2:\$A\$26)

=B2/(СЧЁТ(\$A\$2:\$A\$26)+1)

=НОРМСТОБР(C2)

Статистика в аналитической химии

Ранкит-график первичной обработки данных

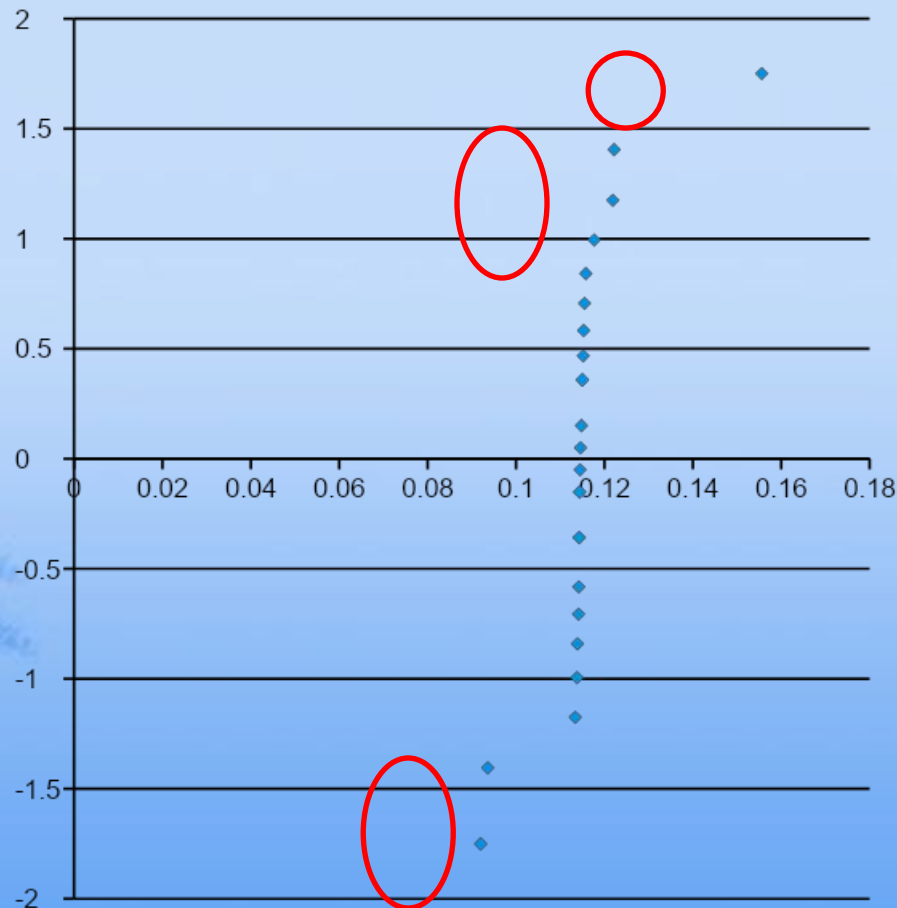


Очевидно, что точка 0.9083 М не является частью нормального распределения, так как она лежит очень далеко от прямой линии, определяющейся другими точками. Можно предположить, что и другие точки также не являются частью нормального распределения, так, что после удаления 0.9083 М построение Ранкит-графика необходимо продолжить.

Статистика в аналитической химии

Повтор ранкит-теста для данных с первым удаленным выпавшим значением

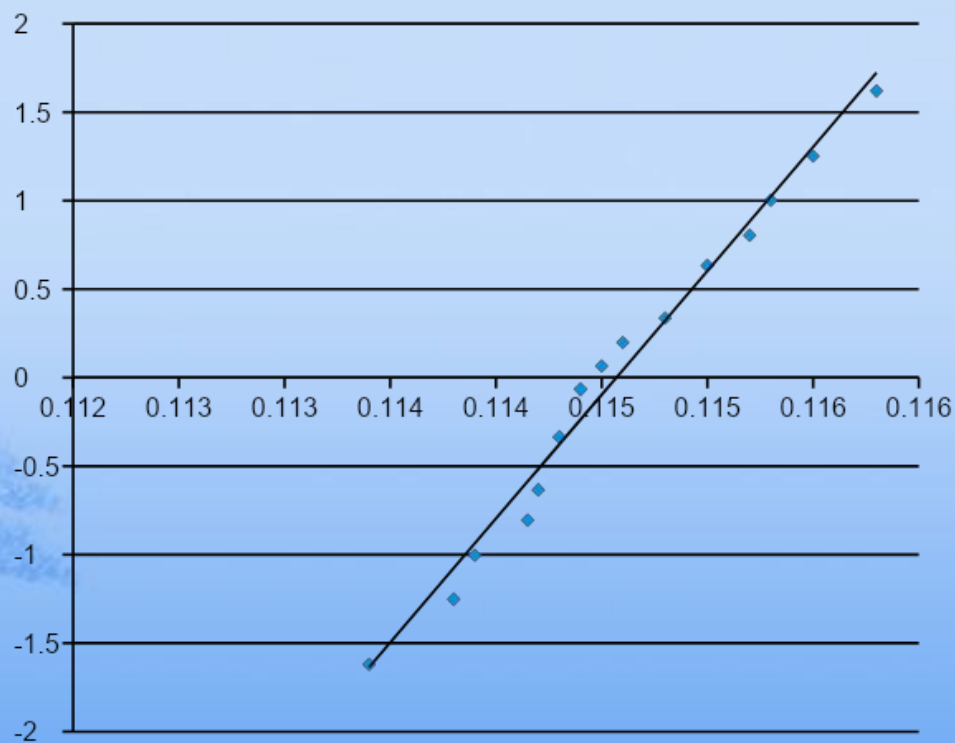
Данные	Кумулятивная частота	Нормализованные значения	z
0.092	1	0.040000	-1.750686071
0.0936	2	0.080000	-1.40507156
0.1134	3	0.120000	-1.174986792
0.1138	4	0.160000	-0.994457883
0.1139	5	0.200000	-0.841621234
0.11415	6	0.240000	-0.706302563
0.1142	7	0.280000	-0.582841507
0.1143	9	0.360000	-0.358458793
0.1143	9	0.360000	-0.358458793
0.1144	11	0.440000	-0.150969215
0.1144	11	0.440000	-0.150969215
0.1145	12	0.480000	-0.050153583
0.1146	13	0.520000	0.050153583
0.1148	14	0.560000	0.150969215
0.115	16	0.640000	0.358458793
0.115	16	0.640000	0.358458793
0.1152	17	0.680000	0.467698799
0.1153	18	0.720000	0.582841507
0.1155	19	0.760000	0.706302563
0.1158	20	0.800000	0.841621234
0.1177	21	0.840000	0.994457883
0.1219	22	0.880000	1.174986792
0.1222	23	0.920000	1.40507156
0.1556	24	0.960000	1.750686071



Статистика в аналитической химии

Завершающий ранжит-тест

Данные	Кумулятивная частота	Нормализованные значения	z
0.1134	1	0.052632	-1.619856259
0.1138	2	0.105263	-1.25211952
0.1139	3	0.157895	-1.003147968
0.11415	4	0.210526	-0.80459638
0.1142	5	0.263158	-0.633640001
0.1143	7	0.368421	-0.33603814
0.1143	7	0.368421	-0.33603814
0.1144	9	0.473684	-0.066011812
0.1144	9	0.473684	-0.066011812
0.1145	10	0.526316	0.066011812
0.1146	11	0.578947	0.199201325
0.1148	12	0.631579	0.33603814
0.115	14	0.736842	0.633640001
0.115	14	0.736842	0.633640001
0.1152	15	0.789474	0.80459638
0.1153	16	0.842105	1.003147968
0.1155	17	0.894737	1.25211952
0.1158	18	0.947368	1.619856259



Статистика в аналитической химии

Очевидно, что точки-данные 0.092, 0.0936, 0.1177, 0.1219, 0.1222, и 0.1556 являются выпавшими значениями. Удаление этих точек и повторное выполнение вычислений дает Ранжит-график, представленный на последнем рисунке, где наконец видно, что точки расположены на прямой линии. В результате только 18 точек являются частью нормального распределения, а 7 выпали из нормального распределения.

Глядя на итоговый график, представленный на рисунке с удаленными всеми выпавшими значениями, можно видеть, что это практически прямая линия. Ваша способность определить линейность на глаз весьма надежна и, вследствие своей простоты этот метод используется наиболее часто.

Статистика в аналитической химии



Тесты на выпавшие значения

Выпавшее значение, есть значение, которое не принадлежит распределению, к которому принадлежат остальные данные. Если включить его в вычисление статистического параметра, такого как среднее или стандартное отклонение, то полученный результат не будет представлять этот параметр.

Поэтому, важно, чтобы выпавшие значения были идентифицированы и исключены из дальнейших вычислений. При повторяющихся химических анализах, выпавшие значения неизбежно появляются на экстремумах (т.е. являются или самым большим, или самым маленьким результатом). Помните, однако, что результаты нельзя просто так отбросить: должна быть основа для идентификации данных как выпавших значений и стратегия работы с ними. Выпавшие значения все же являются результатами и должны исследоваться и включаться в отчет, даже если они не используются последующем анализе. Если данные имеют несколько выпавших значений или содержат данные из двух генеральных совокупностей, то графический метод, такой как Ранкит-график весьма полезен для сортировки нормально распределенных поднаборов результатов. Однако, также полезно иметь быстрый метод для решения вопроса о том, является или нет конкретное значение выпавшим. Метод, который рекомендует ISO является т.н. тест Грубба, хотя многие старые книги предлагают тест Диксона.

Тест Грубба заключается в нахождении критерия G , который потом сравнивается с табличным значением $G_{critical}$ или же это критичное значение рассчитывается по соответствующей формуле.

Если $G > G_{critical}$ то соответствующая точка отбрасывается.

Статистика в аналитической химии

$$G = \frac{|x_{suspect} - \bar{x}|}{s}$$

$$G_{critical} = \frac{(n-1)}{\sqrt{n}} \cdot \sqrt{\frac{t_{(\frac{0.05}{n}, n-2)}^2}{n-2 + t_{(\frac{0.05}{n}, n-2)}^2}}$$

Результат применения Excell для выполнения теста Грубба рассмотрим на следующем примере:

Было проведено определение концентрации кальция в молоке посредством титрования ЭДТА. Было выполнено 10 повторных измерений, в результате которых получены следующие результаты (мг/г) 4.59, 10.00, 6.07, 4.73, 9.91, 5.28, 16.65, 5.17, 4.59 и 4.38. Нужно проверить эту выборку данных на предмет выпадающих значений.

Статистика в аналитической химии

A	B	C
#	Данные	G
1	4,59	0,643013
2	10	0,72279
3	6,07	0,269374
4	4,73	0,607669
5	9,91	0,700068
6	5,28	0,468816
7	16,65	2,401641
8	5,17	0,496587
9	4,59	0,643013
10	4,38	0,696029
Среднее	7,137	
SD	3,961041	
G _{critical}	2,289954	

=ABS(B4-\$B\$13)/\$B\$14

=ABS(B4-\$B\$13)/\$B\$14

=СТАНДОТКЛОН(B2:B11)

=(9/КОРЕНЬ(10))*КОРЕНЬ(СТЮДРАСПОБР
(0,05/10;8)^2/(8+(СТЮДРАСПОБР(0,05/10;8)^2)))

Статистика в аналитической химии

Так как $G_{\text{suspect}} > G_{\text{critical}}$ значение 16.65 мг/г является выпавшим и может быть отвергнуто с уровнем вероятности 95%.

Комментарий

1. Тест Грубба предназначен только для одного выпавшего значения.
2. Отбрасывая выпавшее значение, утверждаем, что оно не является частью нормального распределения. Таким образом, после его исключения можно рассчитать среднее и стандартное отклонение, в нашем случае 6.08 и 2.25, соответственно.
3. Обычный вопрос, который задают: «А после исключения выпавшего значения можно снова проверить самый удаленный от среднего результат?» В идеале тест Грубба предназначен только для одного потенциального выпавшего значения, хотя очень часто приходится видеть, что его применяют и к следующему потенциальному выпавшему значению. Следует быть осторожным, так как если Вы отбрасываете слишком много точек-данных из малого набора данных (скажем из 10 значений), то скорее всего ваши данные не распределены нормально. Для больших наборов данных (больше чем 10 значений) наилучшим способом идентификации выпавших значений является Ранкит-график, где точки, отклоняющиеся от прямой линии не являются частью нормального распределения.

Статистика в аналитической химии



Односторонний t - тест

Систематическая погрешность в аналитической методике должна быть определена и скорректирована. Систематическая погрешность оценивается путем выполнения измерения с привлечением сертифицированного референсного материала (подчас называемом CRM).

Среднее из ряда определений, может быть использовано для решения вопроса является ли систематическая погрешность значимой путем использования одностороннего t – теста, рассчитав критерий Стьюдента по результатам измерений и оценив ассоциированную с этим критерием вероятность.

$$t = \frac{|\mu - \bar{x}| \sqrt{n}}{s} = \frac{|x_{\text{CRM}} - \bar{x}| \sqrt{n}}{s}$$

Для теста на систематическую погрешность нуль гипотеза состоит в том, что систематической погрешности нет.

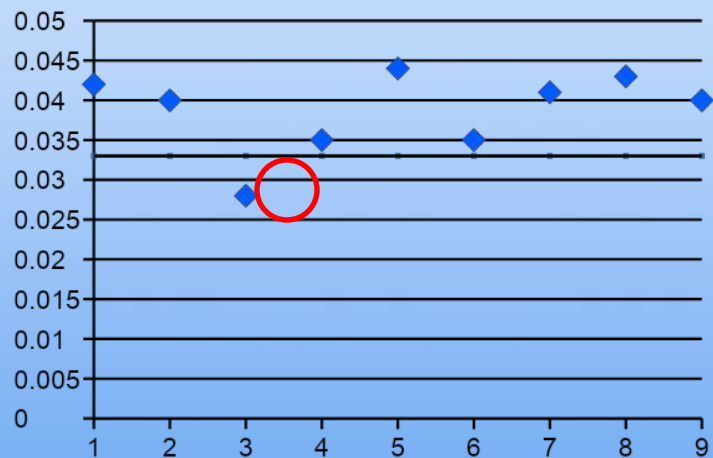
Как и во всех такого рода тестированиях, если вероятность ниже заранее оговоренного предела ($p < \alpha$) то нуль гипотеза (нет систематической погрешности) отвергается и мы заключаем, что таковая погрешность имеется.

Статистика в аналитической химии

ПРИМЕР :

Содержание фтора в зубной пасте согласно метода измеряется с использованием ионселективного электрода. Выполнение измерения требует, чтобы проба была приготовлена путем экстракции фторид-иона из зубной пасты. Для определения того присутствует ли на стадии экстракции или измерения систематическая ошибка ряд аналитиков измеряли содержание фторид-иона в образце пасты с приписанным содержанием 0.033 %. Результаты для 9 аналитиков составили 0.042, 0.040, 0.028, 0.035, 0.044, 0.035, 0.041, 0.043, 0.040%.

Если посмотреть данные, то можно заподозрить наличие систематической ошибки, если все кроме одного значения выше чем приписанное значение, однако, не выполнив тест, нельзя быть уверенным, что это именно так. Построение точек еще более убеждает, что 0.028 % является выпавшим значением. Однако, оценив для этого значения критерий Груббса, получаем значение, говорящее, что полученное значение содержания фторид-иона вписывается внутрь критического диапазона и не может быть отброшено.



Статистика в аналитической химии

A	B	C
Приписное значение содержания фтора в зубной пасте	Результаты, полученные 9 аналитиками	G
0,033	0,042	
	0,04	
	0,028	2,091905
	0,035	
	0,044	
	0,035	
	0,041	
	0,043	
	0,04	
Среднее	0,038666667	
СКО	0,00509902	
G _{critical}	2,215004223	
t _{расчит.}	3,333974297	
P _{ассоц.}	0,010323211	

=ABS(B13-B5)/B14

=СРЗНАЧ(B3:B11)

=СТАНДОТКЛОН(B3:B11)

= (8/КОРЕНЬ(9))*КОРЕНЬ(СТЪЮДРАСПОБР(0,05/9;7)^2/(7+(СТЪЮДРАСПОБР(0,05/9;7)^2)))

= ABS(A3-B13)*КОРЕНЬ(9)/B14

=СТЪЮДРАСП(B17;8;2)

Статистика в аналитической химии

Приписанное значение лежит непосредственно за 95% доверительного интервала, что дает нам возможность предположить, что имеется значимая разница. Заметьте, что нанесение данных на график всегда полезно, но все равно следует вычислять вероятности.

Определив $p = 0.0103$ мы можем заключить, что H_0 принимается с 99% вероятностью.

Но, так как в условии $p < 0.05$ то нуль-гипотеза отвергается на уровне вероятности 95% и делается заключение о том, что имеется систематическая ошибка в методике на основе использования ионоселективного электрода.

Статистика в аналитической химии



Парный t - тест

В некоторых случаях у нас нет такой роскоши, как возможность проводить измерения на одном и том же тест-материале, но мы можем выполнять всего одно измерение для ряда различных тест-материалов двумя методами. Два метода могут быть сравнены путем учета результатов каждой пары одноразовых экспериментов. Если два метода эквивалентны, то математическое ожидание для разницы между результатами, должно тестироваться по отношению к нулю.

В парном t-тесте, таким образом, вычисляются среднее и стандартное отклонение этой разности и затем выполняется расчет t критерия по формуле, приведенной для одностороннего t-теста, но принимая за 0 величину генерального среднего, что имеет место в том случае, когда два метода дают результаты, равные выборочным средним.

$$t = \frac{|\bar{x}_d| \sqrt{n}}{s_d}$$

ПРИМЕР :

Количество кальция в различных образцах молочного порошка (в мг кальция на 1 г молочного порошка) анализировалось двумя методами, один из примененных методов - метод ААС с озолением матрицы, другой представлял собой комплексометрическое титрование.

Статистика в аналитической химии

Тест Материал >	1	2	3	4	5	6	7	8	9
AAS ^a (мг/г)	3.01	2.58	2.582	1.00	1.81	2.83	2.13	5.14	3.20
СТ ^b (мг/г)	2.81	3.20	3.20	3.20	3.35	3.86	3.88	4.13	4.86

^a атомно-абсорбционная спектрофотометрия. ^bКомплексонометрическое титрование.

A	B	C	D
ААС	КТ		Разницы
3,01	2,81		0,2
2,58	3,2		-0,62
2,582	3,2		-0,618
1	3,2		-2,2
1,81	3,35		-1,54
2,83	3,86		-1,03
2,13	3,88		-1,75
5,14	4,13		1,01
3,2	4,86		-1,66
		Среднее	-0,912
		СКО	1,026604
		t - значение	2,665097
		p - значение	0,028579

=A2-B2

=СРЗНАЧ(D2:D10)

=СТАНДОТКЛОН(D2:D10)

=ABS(D12)*КОРЕНЬ(9)/D13

=СТЬЮДРАСП(D15;8;2)

Статистика в аналитической химии

Вероятность того, что между методиками нет разницы, рассчитанная по представленным данным, составляет $(1-0.0286)*100=97.14\%$. Таким образом, мы заключаем, что нет существенной разности между методиками на основе ААС и на основе комплексометрического титрования.

1. Для семи из девяти пар ААС дал меньшее значение, чем комплексометрия, что должно вызвать подозрение аналитика.
2. Без измерений на сертифицированном референсном материале (CRM) мы не можем сказать, какой из двух методов дает результаты с меньшей систематической погрешностью.

Статистика в аналитической химии

Наиболее часто используемые тесты однородности дисперсий двух и более выборок данных

При выборе аналитической методики для анализа проб необходимо учесть ряд вещей, включая требуемый уровень точности. Это делается путем сравнения дисперсий аналитических методик.

Для решения вопроса о том имеется или нет значимая разница между дисперсиями существует несколько критериев. Наиболее популярны критерии Кохрена, Бартлетта, Фишера и как «венец» мощности такого рода исследований – метод ANOVA.

Анализ однородности дисперсий по Кохрену

Назначение. Проверка гипотезы о принадлежности нескольких дисперсий к одной генеральной совокупности. Нулевая гипотеза: $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$

Предпосылки. Данные независимы и распределены по нормальному закону. Количество наблюдений одинаково в каждой выборке. Краткие теоретические сведения. Критериальное значение рассчитывается по следующей формуле:

$$G = \frac{S_{max}^2}{\sum_{i=1}^k S_i^2}$$

где S_{max}^2 — максимальная из дисперсий;
 S_i^2 — эмпирические дисперсии, рассчитанные каждой выборке.

Если $G > G_{\alpha}(k, n)$, то нулевая гипотеза отклоняется. Критические значения можно найти также, пользуясь таблицами F-распределения, с помощью соотношения

Статистика в аналитической химии

$$g_{\alpha}(k, n) = \frac{F_{\frac{k+1-\alpha}{k}}[n-1; (n-1)(k-1)]}{k-1 + F_{\frac{k-1+\alpha}{k}}[n-1; (n-1)(k-1)]}$$

, где $F_{\gamma}(f_1, f_2)$ - γ квантиль F-распределения с f_1 и f_2 степенями свободы.

Пример :

При исследовании комбинированного действия верапамила, тетрадоксина и тетраэтиламония бромида было проведено 9 экспериментов (по 3 раза каждый) по анализу влияния этих веществ на белковый обмен. Необходимо проверить, можно ли считать дисперсии во всех 9 экспериментах однородными. Если они таковыми не являются, это означает, что рассеяние зависит от доз каких-либо из изучаемых препаратов или имеются грубые ошибки в результатах наблюдений, связанные, например, с особенностями используемых в экспериментах животных, и пр.

Для выполнения анализа однородности по Кохрену необходимо выполнить следующие операции:

1. Определить, можно ли считать закон распределения каждой выборки нормальным. Если нет (хотя бы для одной выборки), то следует использовать непараметрический критерий, если да - продолжить.
2. Рассчитать дисперсии по каждому опыту.
3. Найти максимальную из дисперсий.
4. Найти сумму дисперсий.
5. Найти расчетное значение критерия Кохрена.
6. Найти критическое значение критерия Кохрена.
7. Сравнить критерии.

Статистика в аналитической химии

	A	B	C	D	E	F			
1		Анализ однородности дисперсий по Кохрену							
2									
3									
4									
5		Номер	Угнетение синтеза белка, %			Дисперсии			
6			1	2	3				
7									
8		1	30,2	43,6	56,0	166,4933			
9		2	4,0	1,0	3,5	2,5833			
10		3	3,5	2,7	3,5	0,2133			
11		4	31,5	28,7	30,1	1,9600			
12		5	13,6	16,1	21,1	14,5833			
13		6	56,0	28,7	36,2	198,9300			
14		7	16,1	16,1	28,7	52,9200			
15		8	11,1	16,1	13,6	6,2500			
16		9	26,1	35,2	16,0	92,2433			
17									
18		Дисперсия воспроизводимости (среднее)				59,5752			
19		Максимальная дисперсия				198,9300			
20		Сумма дисперсий				536,1767			
21		Число выборок (повторов)				3			
22		Объем выборки				9			
23		Уровень значимости				0,05			
24		Расчетное значение критерия Кохрена				0,37101577			
25		Критическое значение критерия Кохрена				0,43770258			
26									
27		$G < G_{\alpha}$ (0,3715 < 0,4377)							
28									

=ДИСП(C9:E9)

=СРЗНАЧ(F8:F16)

=МАКС(F8:F16)

=СУММ(F8:F16)

=СЧЁТ(C8:E8)

=СЧЁТ(C8:C16)

=F19/F20

=ФРАСПОБР(F23/(F22-1);F21;(F22-1-1)*F21)/(ФРАСПОБР(F23/(F22-1);F21;(F22-1-1)*F21)+F22-1-1)

Статистика в аналитической химии

Анализ однородности дисперсий по Бартлетту

Назначение. Проверка гипотезы о принадлежности нескольких дисперсий к одной генеральной совокупности. Применяется в случае, когда выборки, по которым определяются оценки дисперсий, имеют разный размер (в отличие от критерия Кохрена, требующего выборок равного размера). Нулевая гипотеза : $\sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_k$.

Предпосылки. Распределение данных должно быть нормальным, данные независимы. Данный критерий очень чувствителен к отклонению данных от нормального закона распределения. Если вы сомневаетесь в том, что исследуемые данные распределены в соответствии с нормальным законом распределения, данный критерий лучше НЕ ИСПОЛЬЗОВАТЬ.

Критериальное значение рассчитывается по следующей модифицированной формуле:

$$\chi^2_{kp} = \frac{2}{c} \left(2,3026 \left(v \ln S^2 - \sum_{i=1}^k v_i \ln S_i^2 \right) \right)$$

где $c = \frac{\sum_{i=1}^k \frac{1}{v_i} - \frac{1}{v}}{3(k-1)} + 1$; $S^2 = \frac{\sum_{i=1}^k v_i S_i^2}{v}$

При этом $v = \sum_{i=1}^k v_i$ — общее число степеней свободы; k — число групп; $v_i = n_i - 1$ — число степеней свободы в каждой группе (n_i — объем i -й группы); S_i^2 — оценка дисперсии в каждой группе; S^2 — взвешенная оценка дисперсии.

Статистика в аналитической химии

Если рассчитанное значение χ^2 больше или равно критическому значению, взятому с уровнем значимости α и числом степеней свободы ν , то нулевая гипотеза принимается.

Алгоритм вычислений следующий :

1. Проверить имеющиеся данные (по каждому столбцу отдельно) на соответствие их нормальному распределению. Эту проверку можно выполнить используя критерий Шапиро-Уилка или, например, при помощи функции NORMSAMP_1 (написанной отдельно на VBA, авторы Лапач С.Н, Чубенко А.В., Бабич П.Н.) . Это необходимо сделать **ОБЯЗАТЕЛЬНО**, так как данный критерий очень чувствителен к отклонению анализируемых данных от нормального закона распределения.
2. Вычислить для данных критериальное значение. Использовать формулу, приведенную выше, довольно сложно, поэтому можно применить функцию BARTLET() (написанную отдельно на VBA, авторы см.выше).
3. Для того, чтобы принять или отвергнуть H_0 , необходимо вычислить общее число степеней свободы. Это можно сделать как «вручную», так и при помощи пользовательской функции BARTLET_DF () .

Статистика в аналитической химии

	A	I	J	K	L
1		Анализ однородности дисперсий по Бартлетту			
2					
3					
4					
5		Номер	Угнетение синтеза белка, %		
6			1	2	3
7			1	2	3
8		1	30,2	43,6	56,0
9		2	4,0	1,0	3,5
10		3	3,5	2,7	3,5
11		4	31,5	28,7	30,1
12		5	13,6	16,1	21,1
13		6	56,0	28,7	36,2
14		7	16,1	16,1	28,7
15		8	11,1	16,1	13,6
16		9	26,1	35,2	16,0
17					
18		Проверка на нормальность	NORM	NORM	NORM
19					
20		Общее число степеней свободы V			24
21		Расчетное значение критерия Бартлетта			1,075495
22		Критическое значение критерия Бартлетта			36,41503
23					
24					
25					
26					
27		$\chi^2_{\text{эмп.}} < \chi^2_{\alpha} (1,075 < 36,415)$			
28					

=NORMSAMP_1(K8:K16)

=BARTLET_DF(J8:L16)

=BARTLET(J8:L16)

=ХИ2ОБР(0,05;L20)

Статистика в аналитической химии



ANOVA в Excel

Excel предлагает три варианта применения ANOVA посредством его надстройки Analysis ToolPak («Анализ данных»):
ANOVA: «Однофакторный дисперсионный анализ» ;
ANOVA: «Двухфакторный дисперсионный анализ с повторениями» ; и
ANOVA: «Двухфакторный дисперсионный анализ без повторений» .

Пример посвящен сравнению уровней глюкозы в безалкогольных напитках, определяемых методиками на основе спектрофотометрии и ферментативного электрода (второй альтернативный). Заключение, полученное с использованием t -теста, состояло в том, что эти два средних различны. Аналитическая лаборатория поэтому решила проверить обе методики относительно АОАС (Association of Official Analytical Chemists) который использует ВЭЖХ. Аналитические результаты для шести повторных измерений (единицы mM) следующие:

- Спектрофотометрия: 1.90, 1.82, 1.70, 1.94, 1.85, 1.90
- Ферментативный электрод: 1.35, 1.65, 1.76, 1.41, 1.80, 1.33
- Использование методики АОАС: 1.92, 1.82, 1.85, 1.79, 1.89, 1.95

Эффективным способом решения имеются или нет значительные различия, является однофакторный ANOVA.

Рассмотрим для начала расчеты без применения надстройки «Анализ данных»

Статистика в аналитической химии

	A	B	C	D
1		Спектроф.	Энзимат.	АОАС
2		1.9	1.35	1.92
3		1.82	1.65	1.82
4		1.7	1.76	1.85
5		1.94	1.41	1.79
6		1.85	1.8	1.89
7		1.9	1.33	1.95
8				
9	Среднее	1.852	1.550	1.870
10	SD	0.085	0.212	0.061
11				
12	Grand Mean		1.757	
13				
14		Спектроф.	Энзимат.	АОАС
15		0.143	-0.407	0.163
16		0.063	-0.107	0.063
17		-0.057	0.003	0.093
18		0.183	-0.347	0.033
19		0.093	0.043	0.133
20		0.143	-0.427	0.193

Исходные данные

=СРЗНАЧ(B2:B7)
=СРЗНАЧ(C2:C7)
=СРЗНАЧ(D2:D7)

=СРЗНАЧ(B2:D7)

=B5-\$C\$12
=C5-\$C\$12
=D5-\$C\$12

Статистика в аналитической химии

=6*(CPЗНАЧ(B15:B20)^2)+(6*CPЗНАЧ(C15:C20)^2)+(6*CPЗНАЧ(D15:D20)^2)

=B23/C23

=D23/D25

	A	B	C	D	E	F
22		SS	df	MS	F	P
23	Между факторами	0.3874778	2	0.193739	10.39062	0.001473
24						
25	Внутри фактора	0.2796833	15	0.018646		
26						
27	ИТОГ	0.6671611	17			

=FРАСП(E23;C23;C25)

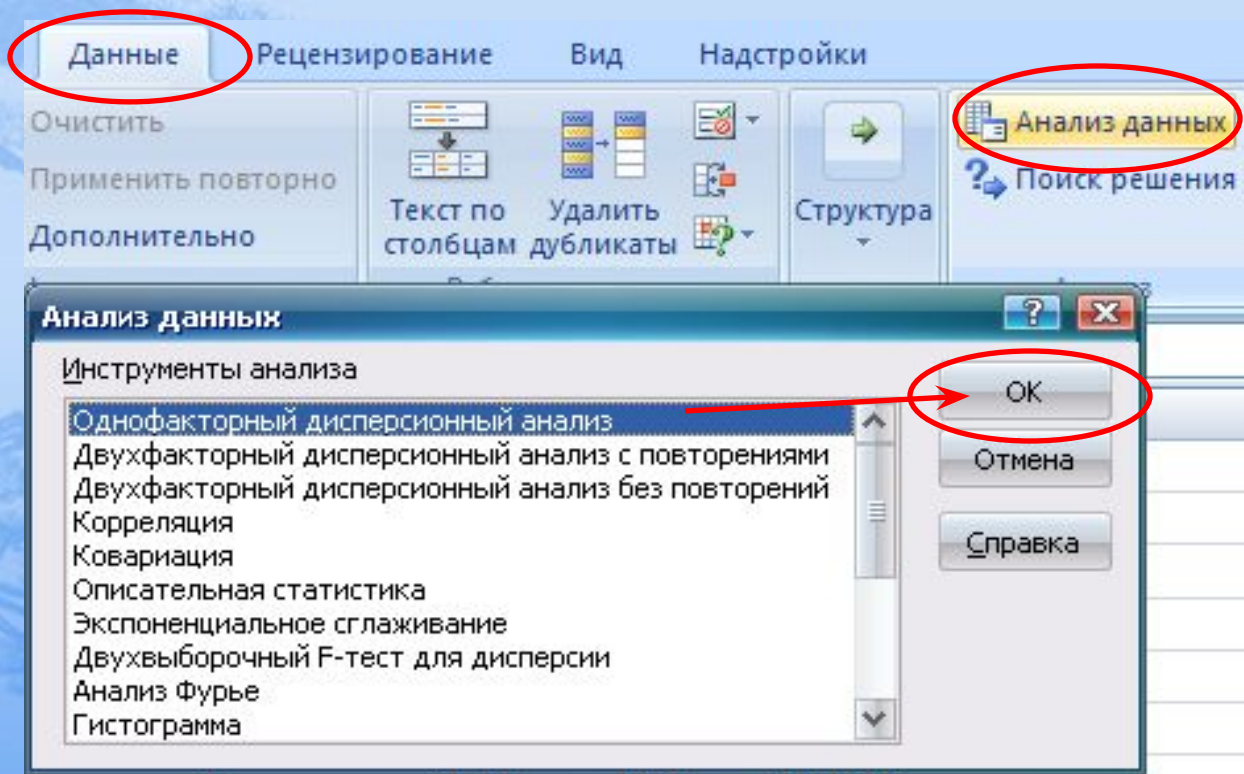
=B27-B23

=B25/C25

=СУММКВ(B15:D20)

Статистика в аналитической химии

С применением надстройки «Анализ данных» все значительно упрощается



Статистика в аналитической химии

	A	B	C
1	Спектроф.	Энзимат.	АОАС
2	1.9	1.35	1.92
3	1.82	1.65	1.82
4	1.7	1.76	1.85
5	1.94	1.41	1.79
6	1.85	1.8	1.89
7	1.9	1.33	1.95
8			

Однофакторный дисперсионный анализ

Входные данные

Входной интервал:

Группирование: по столбцам по строкам

Метки в первой строке

Дельфа:

Параметры вывода

Выходной интервал:

Новый рабочий лист:

Новая рабочая книга

OK
Отмена
Справка

Статистика в аналитической химии

Результат с применением надстройки «Анализ данных»

E	F	G	H	I	J	K
Однофакторный дисперсионный анализ						
ИТОГИ						
<i>Группы</i>	<i>Счет</i>	<i>Сумма</i>	<i>Среднее</i>	<i>Дисперсия</i>		
Спектроф.	6	11.11	1.851667	0.00729667		
Энзимат.	6	9.3	1.55	0.04492		
АОАС	6	11.22	1.87	0.00372		
Дисперсионный анализ						
<i>Источник вариации</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-Значение</i>	<i>F критическое</i>
Между группами	0.387477778	2	0.193739	10.3906203	0.001473225	3.682320344
Внутри групп	0.279683333	15	0.018646			
Итого	0.667161111	17				

Статистика в аналитической химии

Таким образом значение F составляет 10.39062 с ассоциированной вероятностью 0.001473. Так как это меньше 0.05 то это говорит нам, что нуль-гипотеза (отсутствие различий между методами) может быть отвергнута на уровне вероятности 95% ($99.85\% = 100 (1 - 0.001473)$ на самом деле) и, поэтому, мы делаем заключение, что имеется значимая разница в средних связанная с изучаемым фактором, т.е. имеются существенно отличающиеся методики. Т.е. вне связи с тем, как проведены вычисления «вручную» или с использованием встроенного макроса Excel, получается один и тот же ответ.

Статистика в аналитической химии

Реализация метода ANOVA с использованием статистических программных продуктов Minitab16 и Statistica 6.0

Помимо распространенных офисных табличных процессоров типа MS Excel и Sun OpenOffice Calc часто для научных исследований применяют статистические программы из которых следует упомянуть Minitab 16 (LEAD Technologies) и Statistica 6.0 (Statsoft).

Программы очень функциональны и позволяют извлечь из матриц данных максимум статистической информации, зачастую не всегда необходимую. Применение этих программ должно базироваться на принципе соответствия цели. Т.е. если планируется исследование влияния более чем двух факторов, то тогда применение для реализации ANOVA Minitab 16 и Statistica 6.0 оправдано, а если используется лишь одно-двухфакторный анализ, то вполне достаточно MS Excel.

Рассмотрим реализацию ANOVA с помощью Minitab 16 . Эта программа удобна, поскольку позволяет работать с наборами данных как Stacked («сложенных в стопку» в одной колонке) так и Unstacked (расположенных более привычно в виде разных колонок). Воспользуемся набором данных из предыдущего примера по определению концентрации глюкозы спектрофотометрическим методом, энзиматическим и ВЭЖХ(АОАС) методами. Данные вначале расположим Stacked, создав предварительно переменную «Метод», параметрами которой являются повторяющиеся наименования методов.

Статистика в аналитической химии

Minitab - ANOVA_minitab16_primer.MPJ

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Basic Statistics
Regression
ANOVA
One-Way...
One-Way (Unstacked)...
Two-Way...
Analysis of Means...
Balanced ANOVA...
General Linear Model...
Fully Nested ANOVA...
Balanced MANOVA...
General MANOVA...
Test for Equal Variances...
Interval Plot...
Main Effects Plot...
Interactions Plot...

Worksheet 2 ***

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9
	Metod	Data							
1	Spectro	1,90							
2	Spectro	1,82							
3	Spectro	1,70							
4	Spectro	1,94							
5	Spectro	1,85							
6	Spectro	1,90							
7	Enzyme	1,35							
8	Enzyme	1,65							
9	Enzyme	1,76							
10	Enzyme	1,41							
11	Enzyme	1,80							
12	Enzyme	1,33							
13	AOAC	1,92							
14	AOAC	1,82							
15	AOAC	1,85							
16	AOAC	1,79							

One-Way Analysis of Variance

C1 Metod
C2 Data

Response: Data
Factor: Metod

Store residuals
 Store fits

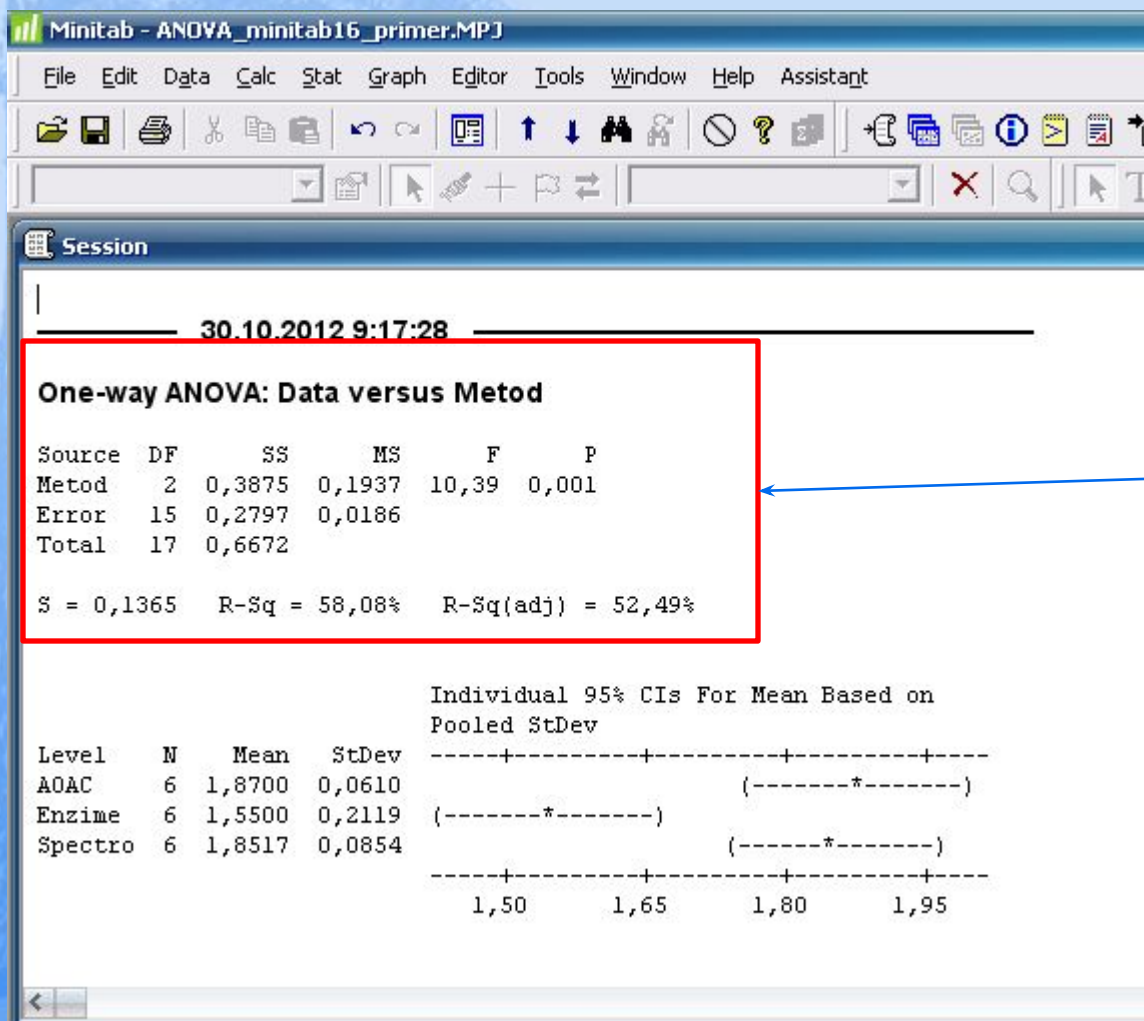
Confidence level: 95,0

Select Comparisons... Graphs...
Help OK Cancel

После выбора вида статистических расчетов
Stat -> ANOVA -> One-Way...

В окне задания параметров One-Way Analysis of Variance
Указывают колонку с данными DATA как Response (отклик),
а в поле Factor (фактор) – METHOD. Указывают уровень
значимости 95 или 99 % и нажимают «Ок».

Статистика в аналитической химии



Результат работы модуля ANOVA будет выведен в окно отчетов «Session».

Статистика в аналитической химии

Minitab - ANOVA_minitab16_primer.MPJ

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Session

30.10.2012 9:17:28

One-Way Analysis of Variance

C1 Спектроф.
C2 Энзимат.
C3 АОАС

Responses (in separate columns):

Store residuals
 Store fits

Confidence level: 95,0

Select Comparisons... Graphs...
Help OK Cancel

Worksheet 3 ***

	C1	C2	C3	C4	C5	C6	C7	C8
	Спектроф.	Энзимат.	АОАС					
1	1,90	1,35	1,92					
2	1,82	1,65	1,82					
3	1,70	1,76	1,85					
4	1,94	1,41	1,79					
5	1,85	1,80	1,89					
6	1,90	1,33	1,95					
7								

Minitab - ANOVA_minitab16_primer.MPJ

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Session

30.10.2012 9:17:28

One-Way Analysis of Variance

C1 Спектроф.
C2 Энзимат.
C3 АОАС

Responses (in separate columns):

'Спектроф.', 'АОАС'

Store residuals
 Store fits

Confidence level: 95,0

Select Comparisons... Graphs...
Help OK Cancel

Worksheet 3 ***

	C1	C2	C3	C4	C5	C6	C7
	Спектроф.	Энзимат.	АОАС				
1	1,90	1,35	1,92				
2	1,82	1,65	1,82				
3	1,70	1,76	1,85				
4	1,94	1,41	1,79				
5	1,85	1,80	1,89				
6	1,90	1,33	1,95				
7							

Статистика в аналитической химии

Minitab - ANOVA_minitab16_primer.MPJ

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Session

30.10.2012 9:17:28

Results for: Worksheet 3

One-way ANOVA: Спектроф.; Энзимат.; АОАС

Source	DF	SS	MS	F	P
Factor	2	0,3875	0,1937	10,39	0,001
Error	15	0,2797	0,0186		
Total	17	0,6672			

S = 0,1365 R-Sq = 58,08% R-Sq(adj) = 52,49%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	-----+-----	
Спектроф.	6	1,8517	0,0854	(-----*-----)	(-----*-----)
Энзимат.	6	1,5500	0,2119	(-----*-----)	(-----*-----)
АОАС	6	1,8700	0,0610	(-----*-----)	(-----*-----)

Worksheet 3 ***

	C1	C2	C3	C4	C5	C6	C7
	Спектроф.	Энзимат.	АОАС				
1	1,90	1,35	1,92				

После выбора вида статистических расчетов Stat -> ANOVA -> One-Way...

В окне задания параметров One-Way Analysis of Variance. Указывают колонки с данными как Responses (отклики) («Спектроф.» – «АОАС»). Указывают уровень значимости 95 или 99 % и нажимают «Ок».

Результат работы модуля ANOVA будет выведен в окно отчетов «Session».

Статистика в аналитической химии

Рассмотрим реализацию ANOVA с помощью Statistica 6.0 . Эта программа работает с наборами данных Stacked («сложенных в стопку» в одной колонке). Воспользуемся набором данных из предыдущего примера по определению концентрации глюкозы спектрофотометрическим методом, энзиматическим и ВЭЖХ(АОАС) методами. Данные расположим Stacked, создав предварительно переменную «Метод», параметрами которой являются повторяющиеся наименования методов как в Minitab 16.



Статистика в аналитической химии

The screenshot shows the STATISTICA software interface. The 'Statistics' menu is open, and 'Analysis of Variance' (Анализ вариантов) is highlighted. A red circle highlights the 'Statistics' menu, and a red arrow points from it to the 'Analysis of Variance' option. Below the menu, a data table is visible with columns '1 Metod' and '2 DATA'.

	1 Metod	2 DATA
1	Spectr	1,9
2	Spectr	1,82
3	Spectr	1,7
4	Spectr	1,94
5	Spectr	1,85
6	Spectr	1,9
7	Enzimat	1,35
8	Enzimat	1,65
9	Enzimat	1,76
10	Enzimat	1,41
11	Enzimat	1,8
12	Enzimat	1,33
13	AOAC	1,92
14	AOAC	1,82
15	AOAC	1,85
16	AOAC	1,79
17	AOAC	1,89
18	AOAC	1,95

The screenshot shows the 'General ANOVA/MANOVA' dialog box. The 'Type of analysis' section has 'One-way ANOVA' selected. The 'Specification method' section has 'Quick specs dialog' selected. The dialog box includes options for 'Weighted moments' and 'DF'.

Quick | OK

Type of analysis:

- One-way ANOVA
- Main effects ANOVA
- Factorial ANOVA
- Repeated measures ANOVA

Specification method:

- Quick specs dialog
- Analysis Wizard
- Analysis syntax editor

Multiple dependent variables can be specified for any type of analysis.

Use One-way ANOVA to analyze designs with a single categorical independent variable (factor).

Weighted moments

DF =

W-1 N-1

Cancel

Options

Open Data

SELECT CASES

Статистика в аналитической химии

The image displays three SPSS dialog boxes for ANOVA analysis, with red circles and arrows highlighting specific elements and Russian text boxes providing instructions.

Top Left Dialog: ANOVA/MANOVA One-Way ANOVA: ANOVA_ZZZ.sta
This dialog has the **Variables** button circled in red. Below it, the **Dependent variables:** field is set to *none*. The **Categorical factor:** field is also *none*. The **Factor codes:** field is *none*. The **Between effect:** field is *none*. Buttons for **OK**, **Cancel**, **Options**, and **Syntax editor** are visible.

Top Right Dialog: Select dependent variables and a categorical predictor (factor):
This dialog shows two lists. The left list contains **1-Metod** and **2-DATA**, with a red box around it containing the text: **Данные, ЛИСТИНГ зависимых переменных**. The right list contains **1-Metod** and **2-DATA**, with a red box around it containing the text: **Факторы, ЛИСТИНГ переменных-факторов**. The **OK** button is circled in red. Below the lists are buttons for **Select All**, **Spread**, and **Zoom**. At the bottom, the **Dependent variable list:** contains the number **2**, and the **Categorical predictor (factor):** contains the number **1**.

Bottom Left Dialog: ANOVA/MANOVA One-Way ANOVA: ANOVA_ZZZ.sta
This dialog shows the **Variables** button circled in red. The **Dependent variables:** field is now **DATA**. The **Categorical factor:** field is **Metod**. The **Factor codes:** button is circled in red. Below it, the **Between effect:** field is **"Metod"**. Buttons for **OK**, **Cancel**, **Options**, and **Syntax editor** are visible.

Bottom Right Dialog: Select codes for indep. vars (factors):
This dialog has the **All** button circled in red. The **Metod:** label is followed by the **All** button and the **Zoom** button. Below, there is a text field containing **"Spectr"-ADAC**. Buttons for **OK** and **Cancel** are also present.

Статистика в аналитической химии

ANOVA Results 1: ANOVA_ZZZ

Profiler Resids Matrix Report
Quick Summary Means Comps

All effects/Graphs
All effects

Alpha values
Confidence limits: .950
Significance level: .050

More results Modify Close Options

STATISTICA - Workbook1* - [Univariate Tests of Significance for DATA (ANOVA_ZZZ.sta)]

File Правка Вид Вставка Формат Статистика Графики Инструменты Данные Workbook Окно Помощь

Arial 10 B I U

Data: ANOVA_ZZZ.sta (2v by 18c)

	1	2			
	Metod	DATA			
1	Spectr	1,9			
2	Spectr	1,82			
3	Spectr				
4	Spectr				
5	Spectr				
6	Spectr				
7	Enzimat				
8	Enzimat				
9	Enzimat				
10	Enzimat				
11	Enzimat				
12	Enzimat				
13	AOAC				
14	AOAC				

Workbook1* - Univariate Tests of Significance for DATA (ANOVA_ZZZ.sta)

Univariate Tests of Significance for DATA (ANOVA_ZZZ.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

Effect	SS	Degr. of Freedom	MS	F	p
Intercept	55,58094	1	55,58094	2980,922	0,000000
Metod	0,38748	2	0,19374	10,391	0,001473
Error	0,27968	15	0,01865		

Статистика в аналитической химии

Критические значения статистики W-критерия Вилкоксона

* это непараметрический аналог парного критерия Стьюдента (t-критерий для зависимых выборок)

Критические значения статистики W-критерия Вилкоксона широко применяют при использовании методов проверки статистических гипотез, основанных на рангах. **Ранговые критерии** — это статистические тесты, в которых вместо выборочных значений используются их ранги (номера элементов в упорядоченной по возрастанию выборке). Большинство ранговых критериев являются непараметрическими. Данные о критических значениях достаточно хорошо табулированы — их приводят в виде таблиц во многих литературных источниках. Но необходимая книга не всегда есть под рукой. Поэтому, предлагается использовать для расчета критических значений статистики W формулу, основанную на ее аппроксимации с использованием нормального распределения.

Для конкретного набора значений количества уровней рангов легко построить функцию распределения Вилкоксона. При этом считается, что вероятность любого сочетания пар рангов двух экспертов равновероятна.

Нижнюю критическую точку статистики W довольно точно вычисляют по аппроксимационной формуле:

$$w(Q; m; n)^* = \left[\frac{m(m+n+1) - 1}{2} - \Psi(1-Q) \sqrt{\frac{mn(m+n+1)}{12}} \right]$$

где $[]$ — операция округления к ближайшему целому; Q — уровень значимости для одностороннего критерия; m, n — размеры исследуемых выборок, причем $m \leq n$, $\Psi(1-Q)$ значение обратной функции нормального распределения с параметрами $(0, 1)$. Приведенная выше аппроксимационная формула дает удовлетворительные результаты при $n \geq m \geq 5$.

Верхняя критическая точка связана с нижней соотношением: $W(Q; m; n) = m(m+n+1) - w(Q; m; n)$

Статистика в аналитической химии

Чтобы автоматизировать вычисление критических значений статистики W , необходимо определить свою пользовательскую функцию, например, $W_{perculc}(alfa; m_size; n_size)$ при помощи встроенного в Microsoft Excel языка программирования Visual Basic (написанной авторами Лапач С.Н., Чубенко А.В., Бабич П.Н.).

	A	B	C	D
1		Пример расчета критических значений статистики W ранговых сумм Вилкоксона		
2		Вилкоксона		
3		Уровень значимости		0,05
4		Объем выборки m		8
5		Объем выборки n		10
6				
7		Нижнее критическое значение		
8				
9		Вычисление по формуле		57
10		Вычисление по функции		57
11				
12		Вычисление верхнего критического значения W		95
13				

=ОКРУГЛ((D4*(D4+D5+1)-1)/2-НОРМСТОБР(1-D3)*КОРЕНЬ(D4*D5*(D4+D5+1)/12); 0)

=Wperculc(D3;D4;D5)

=D4*(D4+D5+1)-D9

Статистика в аналитической химии

Двухвыборочный критерий Вилкоксона

Назначение. Проверка гипотезы о равенстве средних двух независимых выборок.

Нулевая гипотеза. Обе выборки имеют одинаковое распределение, то есть извлечены из одной генеральной совокупности. Следствием чего является равенство двух средних.

Предпосылки. Все случайные величины взаимно независимы.

1. При помощи функции `NORMSAMP_1 ()` проверить имеющиеся данные (по каждому столбцу отдельно) на соответствие (или не соответствие) их нормальному распределению.
2. Определить размеры обеих выборок, а также уровень значимости α (0,05).
3. Вычислить верхнее и нижнее критические значения статистики W , используя формулы или пользовательскую функцию, `Wperculcl (alfa; m_size; n_size)`.
4. Рассчитать критериальное значение $W_{эмп}$. Для упрощения этого процесса можно использовать пользовательскую функцию `W_crit1(R_1; R_2)`,

Статистика в аналитической химии

	G	H	I	J
1	Применение двухвыборочного критерия Вилкоксона			
2				
3			Контроль	Эксперимент
4			27	55
5			41	24
6			34	40
7			44	60
8			69	39
9			38	28
10			24	22
11			30	37
12			37	72
13			59	42
14			43	33
15			39	41
16				25
17				
18	Проверка на нормальность	NO_NORM	NO_NORM	
19				
20	Объем выборки 1 N_1			12
21	Объем выборки 2 N_2			13
22	Уровень значимости α			0,05
23	W крит. нижняя граница			119
24	$W_{эмп}$			161
25	W крит. верхняя граница			193

=NORMSAMP_1(J4:J16)

=СЧЁТ(I4:I16)

=СЧЁТ(J4:J16)

=Wperculc(J22/2;J20;J21)

=W_crit1(I4:I15;J4:J16)

=J20*(J20+J21+1)-J23

Статистика в аналитической химии

Критерий Т- Вилкоксона для парных наблюдений (одновыборочный критерий Вилкоксона, критерий знаковых рангов)

Назначение. Аналог t-критерия Стьюдента для парных наблюдений в случае нечисловых данных или закона распределения, отличного от нормального. Применяют для связанных пар наблюдений (проверка отсутствия эффекта обработки).

Нулевая гипотеза. Распределение разностей пар симметрично относительно нуля.

Предпосылки. Разности независимы, принадлежат непрерывной совокупности (необязательно одной и той же), симметричной относительно нуля.

1. Рассчитать значения разниц пар двух выборок.
2. Проранжировать абсолютные значения разниц пар в возрастающем порядке.
3. Назначить им соответствующие значения рангов от 1 и выше. Нулевые значения разницы не рассматриваются (отсюда — уменьшение общего количества наблюдений N). В общем, построение рангов выполняют как обычно, только знаки значений разниц при этом не учитывают).
4. Знаком ранга считать знак соответствующей ему разниц.
5. Посчитать сумму R положительных рангов.
6. Вычислить критериальное значение:

$$T^+ = \sum_{i=1}^n R_i^+$$

$$T = \frac{R - N(N+1)}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

7. Проверить критерий.

Для малых выборок ($N < 20$), если значение R больше табличного значения критерия знаковых рангов Вилкоксона T^* (не следует путать со статистикой ранговых сумм W!) с уровнем значимости $\alpha/2$ и числом степеней свободы N, нулевая гипотеза отвергается.

* Критерий применяется для сопоставления показателей, измеренных в двух разных условиях на одной и той же выборке. Он позволяет установить не только направленность изменений, но и их выраженность.

Статистика в аналитической химии

Примечание. Данный критерий устойчив к умеренным отклонениям от принятых предпосылок. Критерий Вилкоксона не следует применять к нормально распределенным совокупностям. Лучше использовать более мощный t-критерий Стьюдента. Критические значения можно рассчитать по аппроксимирующей формуле :

$$T^+(\alpha) = \left[\frac{n(n+1)}{4} + \sqrt{\frac{n(n+1)(2n+1)}{24}} u_\alpha \right]$$

$$u_{0,95} = 1,645$$

Статистика в аналитической химии

=L4-M4

=ABS(O4)

=Rank1(Q4;\$Q\$4:\$Q\$13;1)

=ЕСЛИ(ЗНАК(O9)=1;R9;0)

=ЧСТРОК(L4:L13)

=СУММ(S4:S13)

=(S16-S15*(S15+1)/4)/(КОРЕНЬ(S15*(S15+1)*(2*S15+1)/24))

=1-НОРМСТРАСП(S17)

=ОКРУГЛ((S15*(S15+1)/4) + КОРЕНЬ
((S15*(S15+1)*(2*S15+1))/24)*1,645;0)

	L	M	N	O	P	Q	R	S	T
1	Применение одновыборочного Т-критерия Вилкоксона								
2									
3	1-я выб.	2-я выб.	Разница	Абс.знач.	Ранги	Полож. ранги			
4	27	20	7	7	5,5	5,5			
5	40	31	9	9	8	8			
6	24	19	5	5	4	4			
7	35	37	-2	2	2	0			
8	60	52	8	8	7	7			
9	70	50	20	20	10	10			
10	27	24	3	3	3	3			
11	26	25	1	1	1	1			
12	39	27	12	12	9	9			
13	38	31	7	7	5,5	5,5			
14									
15	Число экспериментов N						10		
16	Сумма положительных рангов R^+ , T^+						53		
17	Критериальное значение верхнего хвоста знаковых рангов Вилкоксона T^*						2,599200676		
18	Расчетное значение уровня значимости						0,004672057		
19	Критические значения T^+ (α) одновыборочного критерия Вилкоксона						44		
20									
21									

Статистика в аналитической химии



Калибровка и ее параметры с помощью Excel

Калибровка является сердцем химического анализа и процесса посредством которого отклик прибора (в метрологии называется «показание измерительного прибора») связано с значением измеряемой величины, в химии этой величиной чаще всего является концентрация аналита в испытуемом или стандартном (стандартных) растворах. Без надлежащей калибровки измерения прибора не являются прослеживаемыми (traceable) и даже просто корректными.

Проведя лучшую прямую линию через точки можно нанести на линии значение измеренного сигнала для испытуемого раствора и опустив перпендикуляр на ось концентраций найти его концентрацию. В настоящее время по калибровочным данным рассчитывается т.н. регрессионное уравнение, которое затем используется для получения концентрации испытуемого раствора. Хотя и нет более абсолютной необходимости построения графика для определения параметров калибровочной прямой, но хорошей практикой является построение графика в качестве быстрой визуальной проверки на наличие выпавших значений или на возможное наличие кривизны. Поскольку, что мы можем выбирать значения калибровочных концентраций, то концентрация является независимой переменной, а вот отклик прибора является зависимой переменной (потому, что отклик прибора зависит от концентрации аналита).

Самая используемая модель калибровки - линейная $Y = a + bx$, где Y показание прибора, x независимая переменная, которая в большинстве случаев представляет собой концентрацию аналита в растворе, a и b коэффициенты модели известные как «пересечение (intercept)» и «наклон (slope)» и которые определяются из ряда измерений для при различных значениях.

Статистика в аналитической химии

Любое конкретное измерение (y_i) будет подвержено погрешностям измерения, таким образом:

$$Y_i = a + bx_j + \varepsilon_i$$

Аналитические методы, для которых может быть проведено вычитание соответствующего измерения для холостой пробы принуждают калибровочную прямую проходить через нуль.

$$Y = bx$$

Даже для простого линейного уравнения при принятии различных допущений о данных мы приходим к различным значениям **a** и **b**. Большинство электронных таблиц и калькуляторов выполняют «классическую» линейную регрессию. Допущения следующие :

1. Линейная модель корректна (т.е., отклик измерительного прибора на самом деле линейно связан с концентрацией).
2. Все неопределенности заключаются в зависимой переменной (**Y**) и распределены нормально.
3. Данные имеют гомоскедастический характер, это означает, что погрешности в **y** не зависят от концентрации. Данные для которых неопределенность, например, растет с концентрацией называются гетероскедастическими.

Большинство аналитических систем нарушают то или иное из этих допущений, но о разумной линейности обычно можно не беспокоиться.

Регрессия будет избегать очень больших отклонений за счет малых и таким образом, единственная «порочная» точка может исказить всю калибровку. Склонность точки тянуть линию на себя на сленге известна как «принцип рычага».

Коэффициенты уравнений вычисляются из данных посредством следующих формул:

Статистика в аналитической химии

$$b = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

$$a = \bar{y} - b \cdot \bar{x}$$

Эти коэффициенты оценивают действительную функцию, ограниченную неминуемым разбросом, который возникает во время измерения. Точность расчета количественно выражают остаточным средним квадратичным отклонением, S_y , которое является мерой разброса значений измеряемых величин, вокруг градуировочной линии, в соответствии с уравнением :

$$S_y = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N-2}} = \sqrt{\frac{\sum_{i=1}^N [y_i - (a + b \cdot x_i)]^2}{N-2}}$$

Статистика в аналитической химии

Стандартные отклонения могут быть вычислены и для коэффициентов :

$$S_b = \frac{S_y}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2}} \quad S_a = S_y \sqrt{\frac{\sum_{i=1}^N x_i^2}{N \cdot \sum_{i=1}^N (x_i - \bar{x})^2}}$$

Стандартное отклонение определяемой концентрации из m измерений испытуемого раствора, давшего средний отклик y_0 равно :

$$S_{\hat{x}_0} = \frac{S_y}{b} \cdot \sqrt{\frac{1}{m} + \frac{1}{N} + \frac{(y_0 - \bar{y})^2}{b^2 \cdot \sum_{i=1}^N (x_i - \bar{x})^2}}$$

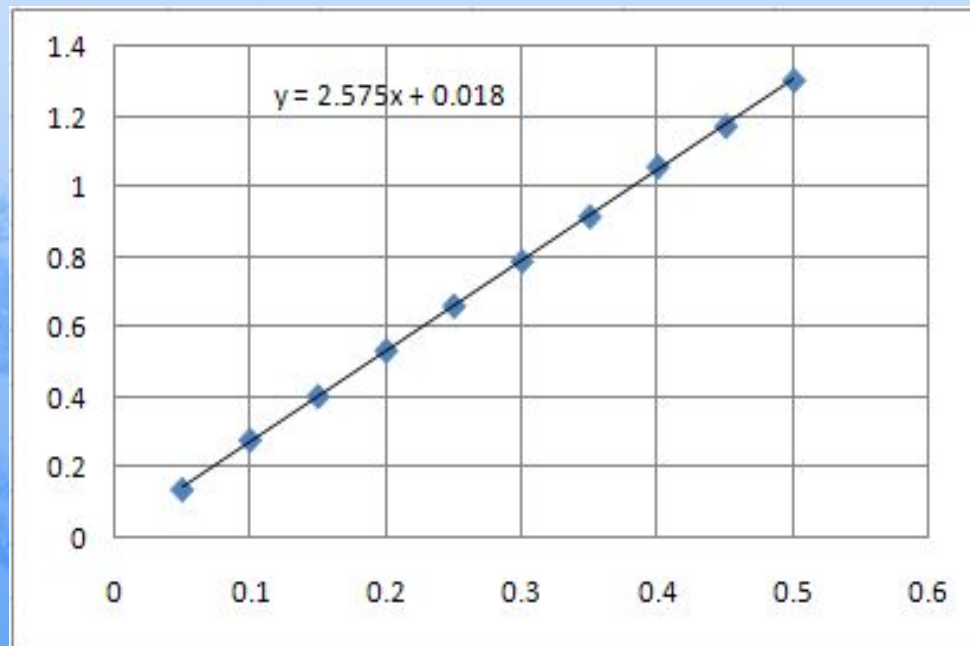
Статистика в аналитической химии

Рассмотрим возможность реализации вышеприведенных расчетов в Excel на примере определения содержания нитрит-иона в воде.

Данные для построения калибровки

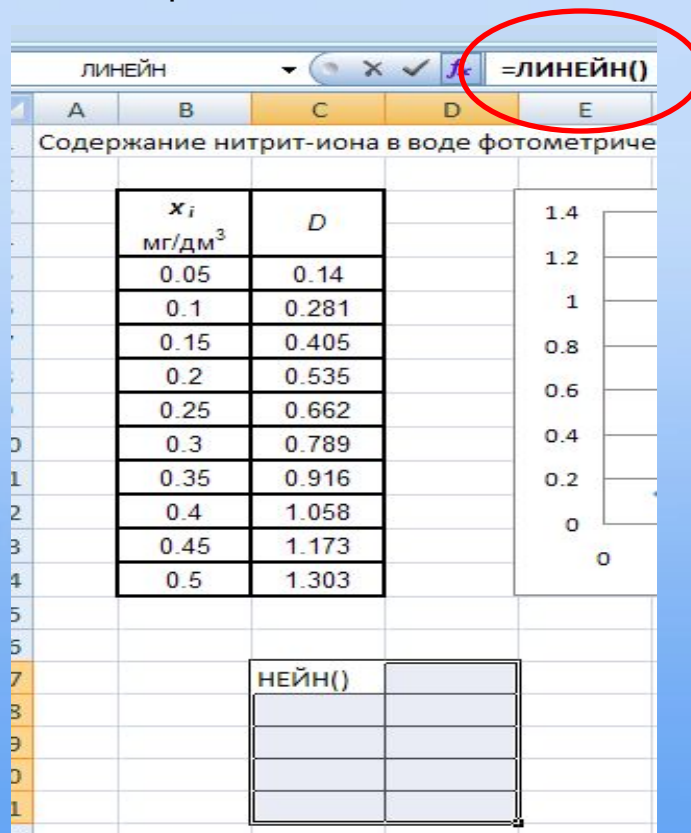
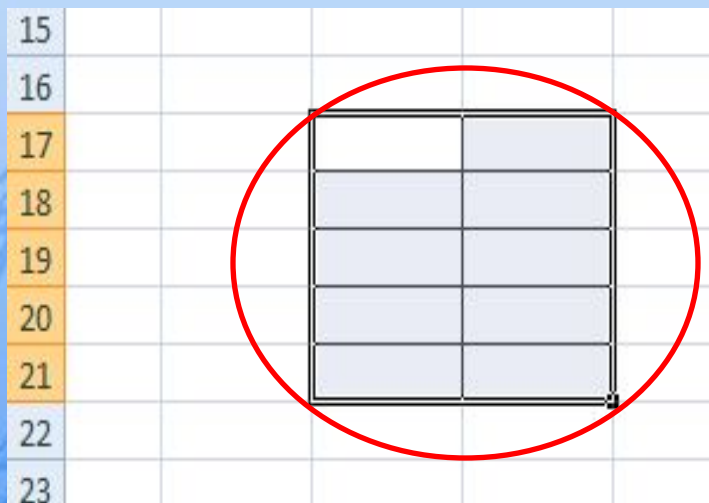
	В	С
1	x_i	D
2	мг/дм ³	
3	0.05	0.14
4	0.1	0.281
5	0.15	0.405
6	0.2	0.535
7	0.25	0.662
8	0.3	0.789
9	0.35	0.916
10	0.4	1.058
11	0.45	1.173
12	0.5	1.303

С помощью мастера диаграмм легко вывести графическое изображение и линию тренда



Статистика в аналитической химии

Для реализации вышеприведенных расчетов в Excel необходимо воспользоваться функцией-массивом – «ЛИНЕЙН» из набора статистических функций данного табличного процессора. Для корректного вывода необходимой информации необходимо в свободной от данных зоне выделить локальный диапазон из двух колонок и пяти строк и выполнить в командной строке вызов функции «ЛИНЕЙН».



Статистика в аналитической химии

После вызова функции «ЛИНЕЙН» появится окно задания необходимых входных диапазонов данных

The screenshot shows the Microsoft Excel interface with the 'ЛИНЕЙН' (LINEAR) function dialog box open. The dialog box is titled 'Аргументы функции' (Function Arguments) and contains the following fields:

- Известные_значения_y** (Known_y_values): C5:C14
- Известные_значения_x** (Known_x_values): B5:B14
- Конст** (Constant): 1
- Статистика** (Statistics): 1

The dialog box also displays the following values and descriptions:

- Известные_значения_y**: = {0.14;0.281;0.405;0.535;0.662;0.789;0.916;1.058;1.173;1.303}
- Известные_значения_x**: = {0.05;0.1;0.15;0.2;0.25;0.3;0.35;0.4;0.45;0.5}
- Конст**: = ИСТИНА
- Статистика**: = ИСТИНА
- Значение**: 2.575272727
- Возвращает**: Возвращает параметры линейного приближения по методу наименьших квадратов.
- Статистика**: логическое значение, которое указывает, требуется ли вернуть дополнительную статистику по регрессии (ИСТИНА) или только коэффициенты m и константу b (ЛОЖЬ или отсутствие значения).

The background shows a scatter plot of the data points from the table below, with a red arrow pointing to the data point at x=0.25, y=0.662. The formula bar shows the function: =ЛИНЕЙН(C5:C14;B5:B14;1;1).

x_i мг/дм ³	D
0.05	0.14
0.1	0.281
0.15	0.405
0.2	0.535
0.25	0.662
0.3	0.789
0.35	0.916
0.4	1.058
0.45	1.173
0.5	1.303

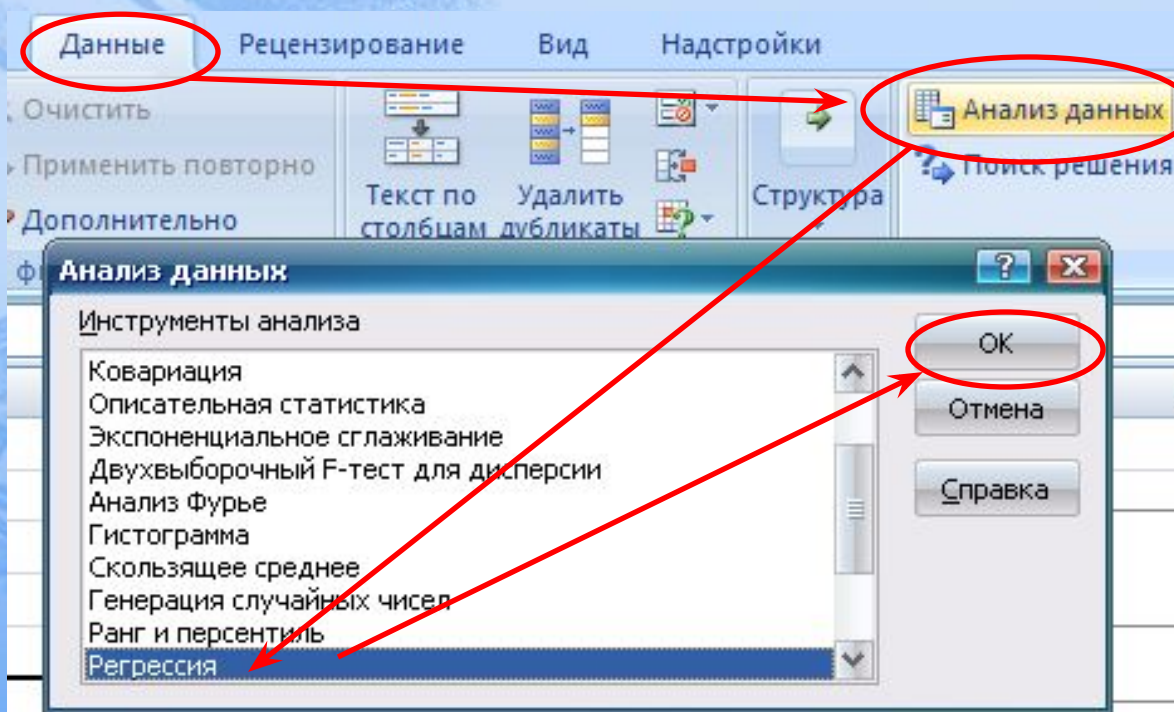
Статистика в аналитической химии

Соответствие полученных данных функции «ЛИНЕЙН» параметрам линейной регрессии

b	a
S_b	S_a
r^2	S_y
Критерий Фишера F	df
$SS_{\text{residual}} = df \cdot S_y^2$	S_y^2

Статистика в аналитической химии

Получение расширенных параметров линейной регрессии, согласно ISO 11095:1996 «Linear calibration using reference materials» с помощью пакета анализа



Статистика в аналитической химии

	A	B	C	D	E	F	G	H	I	J	K
1											
2		x_i	D								
3		мг/дм ³									
4		0,05	0,14								
5		0,1	0,281								
6		0,15	0,405								
7		0,2	0,535								
8		0,25	0,662								
9		0,3	0,789								
10		0,35	0,916								
11		0,4	1,058								
12		0,45	1,173								
13		0,5	1,303								
14											
15											
16											
17											
18											

Регрессия

Входные данные

Входной интервал Y:

Входной интервал X:

Метки

Константа - ноль

Уровень надежности: %

Параметры вывода

Выходной интервал:

Новый рабочий лист:

Новая рабочая книга

Остатки

Остатки

График остатков

Стандартизованные остатки

График подбора

Нормальная вероятность

График нормальной вероятности

OK

Отмена

Справка

Статистика в аналитической химии

A	B	C	D	E	F	G	H	I	J	K
	x_i мг/дм ³	D								
	0,05	0,14								
	0,1	0,281								
	0,15	0,405								
	0,2	0,535								
	0,25	0,662								
	0,3	0,789								
	0,35	0,916								
	0,4	1,058								
	0,45	1,173								
	0,5	1,303								
				Вывод итогов						
				Регрессионная статистика						
				Множественный R	0,999921971					
				R-квадрат	0,999843947					
				Нормированный R-квадрат	0,999824441					
				Стандартная ошибка	0,005165885					
				Наблюдения	10					
				Дисперсионный анализ						
					<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>	
				Регрессия	1	1,367856109	1,367856109	51256,744	1,6217E-16	
				Остаток	8	0,000213491	2,66864E-05			
				Итого	9	1,3680696				
					<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
				Y-пересечение	0,018	0,003528971	5,100636975	0,000929119	0,009862178	0,026137822
				Переменная X 1	2,575272727	0,011374904	226,399523	1,6217E-16	2,549042153	2,601503302
				Вывод остатка						
				<i>Наблюдение</i>	<i>Предсказанное Y</i>	<i>Остатки</i>				
				1	0,146763636	-0,006763636				
				2	0,275527273	0,005472727				
				3	0,404290909	0,000709091				
				4	0,533054545	0,001945455				
				5	0,661818182	0,000181818				
				6	0,790581818	-0,001581818				



Статистика в аналитической химии

Метод контроля стабильности калибровки, согласно ГОСТ Р ИСО 11095-2007

Если функция калибровки должна быть использована в течение большого периода времени, желательно применять метод контроля для проверки справедливости функции калибровки, а также для идентификации, а затем устранения источников нежелательных отклонений. Метод позволяет контролировать на регулярной основе измерительную систему и быстро обнаружить беспорядочное поведение или изменения измерительной системы, в результате которых применение имеющейся функции калибровки становится бесполезным или даже вредным.

Это достигается на основе мониторинга результатов измерений (после преобразований в соответствии с функцией калибровки) на m RM (reference material - образец сравнения) с помощью контрольной карты.

Контрольную карту сначала строят на основе данных, собранных в процессе эксперимента по калибровке. Затем контрольную карту используют для решения о необходимости повторного построения функции калибровки. Ту же самую контрольную карту используют для оценки неопределенности результатов измерений после их преобразования в соответствии с функцией калибровки.

Рассчитывают контрольные границы :

$$U_c = \frac{S_y}{b} \cdot t_{(0,05/2;(NK-2))}$$
$$L_c = -\frac{S_y}{b} \cdot t_{(0,05/2;(NK-2))}$$

Статистика в аналитической химии

Отбирают m РМ так, чтобы соответствующие им принятые значения заполняли диапазон значений, характерных для нормальных режимов работы измерительной системы. Необходимо не менее двух РМ. Рекомендуется три РМ. Предпочтительно (но необязательно) использовать РМ, которые отличаются от используемых в процессе эксперимента по калибровке. Регулярно (например, один раз в день или один раз при каждом изменении) выполняют одно измерение каждого из этих РМ.

Находят преобразованные значения x ($i = 1 \dots t$) для каждого из m РМ. Вычисляют разности между преобразованными значениями x_i^* и принятыми значениями x , передаваемыми РМ.

Откладывают на контрольной карте полученные контрольные значения в соответствии с моментами времени, в которые были проведены измерения m РМ.

Решение о состоянии системы

Если одно или несколько значений, попадают вне контрольных границ это означает, что система находится в неуправляемом состоянии. Необходимо повторить измерения на РМ. Если хотя бы один из новых результатов измерений m РМ попадает в область вне контрольных границ, в этой точке необходимо провести исследование причин такой ситуации. В зависимости от особенностей задачи может быть принято решение о необходимости повторного определения функции калибровки с новым экспериментом по калибровке.

Карты Шухарта

Контрольные карты — график изменения параметров выборки, обычно средних и среднеквадратичного отклонения. Различают контрольные карты для количественного и качественного признаков. Цель построения контрольной карты — выявление точек выхода процесса из устойчивого состояния для последующего установления причин отклонения и его устранения.

Задачи построения КК:

- определить возможности процесса,
- определить точки флуктуации,
- спрогнозировать качество процесса.

Контрольные карты были предложены Шухартом (Shewhart) еще в 20-е годы, в период его работы в телефонной компании Bell Laboratories (США), поэтому их иногда называют карты Шухарта. С помощью этих карт выявляется различие между изменчивостью технологического процесса, вызванной простыми причинами, и изменчивостью, появившейся под действием неслучайных причин.

Карты Шухарта

Контрольные карты используются для оценки "контролируемости" или "неконтролируемости" процесса. Эту оценку можно получить:

- осуществляя проверку замеров важнейших параметров, например, веса сахара в одной упаковке, диаметра отверстия, просверленного в листе металла, коэффициента возврата Recovery;
- осуществляя проверку отдельных качественных характеристик изделия, например, прочно ли упакован пакет, правильно ли закрыта крышкой бутылка, не повреждено ли фарфоровое изделие и т.д.

В первом случае используются контрольные карты количественного признака, а во втором — контрольные карты качественного признака. Существует несколько видов карт Шухарта. Мы остановимся более подробно на двух типах: контрольной карте средних арифметических и контрольной карте индивидуальных значений.

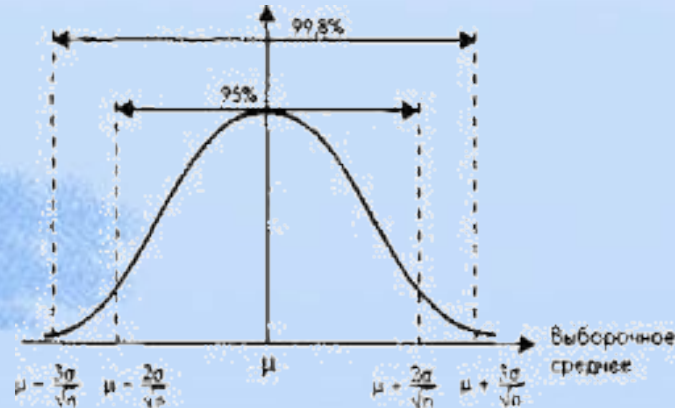
Карты Шухарта

КОНТРОЛЬНАЯ КАРТА СРЕДНИХ АРИФМЕТИЧЕСКИХ

Если генеральная совокупность имеет нормальное (или близкое к нормальному) распределение со средним значением μ и стандартным отклонением s , выборочное распределение выборочного среднего также является нормальным и имеет такое же среднее значение и стандартную ошибку, равную $\frac{s}{\sqrt{n}}$, где n — объем выборки. Для любого нормального распределения между граничными значениями, равными $\mu \pm 2 \cdot S$, заключено примерно 95% распределения. Вероятность того, что полученное значение окажется больше, чем $\mu + 2 \cdot S$, составляет 2,5%, или один случай из 40, вероятность получения значения, меньшего $\mu - 2 \cdot S$, также составляет 2,5%. Аналогично интервал $\mu \pm 3 \cdot S$ охватывает около 99,8% распределения. Вероятность того, что полученное значение превысит $\mu + 3 \cdot S$ или окажется меньше, чем $\mu - 3 \cdot S$, составляет 0,1%, т.е. это событие будет иметь место в одном случае из 1000.

Графическая иллюстрация этих крайних значений приведена на рис. для выборочного распределения выборочного среднего. Эта диаграмма является основой для составления контрольной карты среднего арифметического процесса.

Карты Шухарта



Для построения графика, приведенного на рис., необходимо, чтобы значения μ и σ были известны. Их оценки получают по результатам расчетов среднего значения и стандартного отклонения соответствующих параметров процесса на протяжении длительного промежутка времени. 95%-ные границы распределения называются **верхней и нижней предупреждающими границами (ВПГ и НПГ)**. 98%-ные границы распределения называются **верхней и нижней корректировочными границами (ВКГ и НКГ)**.

Построение контрольной карты состоит в нанесении на график выборочных средних в соответствии с номером выборки.

Стандартная процедура использования этих контрольных карт состоит из следующих шагов:

Карты Шухарта

1. Через равные промежутки времени проводится выборка объемом n и рассчитывается выборочное среднее.
2. Полученное значение выборочного среднего наносится на контрольную карту в соответствии с номером выборки.
3. Если выборочное среднее лежит за пределами границы корректирования, производится остановка процесса в целях выявления неслучайных причин вариации.
4. Если два последовательно полученных значения выборочных средних находятся в промежутке между предупреждающей границей и границей регулирования, предпринимаются немедленные действия по остановке процесса производства и выявлению неисправностей. Если некоторое среднее значение лежит за пределами предупреждающих границ, следующая выборка производится сразу же, до момента проведения очередной выборки.
5. Если точки на графике образуют явный возрастающий или убывающий тренд, предпринимаются определенные меры даже в случаях, когда эти точки находятся в пределах предупреждающих границ. Этот тренд может оказаться индикатором наличия неслучайных причин.

Карты Шухарта

Процедура ведения КК в сущности представляет собой не что иное, как проверку гипотез. Нулевая гипотеза заключается в том, что технологический процесс находится под контролем, причем все технологические параметры соответствуют установленным производственным возможностям. Альтернативная гипотеза утверждает, что процесс не является контролируемым. Каждый раз, когда мы проводим выборку, осуществляется процедура проверки гипотез. Если выборочное среднее лежит за предупреждающими границами, H_0 отклоняется при 5%-ном уровне значимости. Если оно находится за пределами границ регулирования, мы отклоняем H_0 при 2%-ном уровне значимости. С помощью контрольных карт можно без труда показать результаты периодически повторяющейся проверки гипотез.

Карты Шухарта

Формулы расчета средней линии и границ для карт средних

Стати с-тика	Стандартные значения не задано		Стандартные значения задано		
	Средняя линия	ВКГ и НКГ	Средняя линия	ВКГ	НКГ
\bar{X}	$\bar{\bar{X}}$	$\bar{\bar{X}} \pm A_2 \cdot R$ или $\bar{\bar{X}} \pm A_3 \cdot S$	$X_0 (\mu)$	$X_0 \pm A \cdot \sigma_0$	
R	\bar{R}	$D_3 \cdot \bar{R}$ $D_4 \cdot \bar{R}$	R_0 ($d_2 \sigma_0$)	$D_1 \sigma_0$	$D_2 \sigma_0$
S	\bar{S}	$B_3 \cdot \bar{S}$ $B_4 \cdot \bar{S}$	S_0 ($C_4 \sigma_0$)	$B_5 \sigma_0$	$B_6 \sigma_0$

Карты Шухарта

Контрольные карты для индивидуальных значений (X)

В некоторых ситуациях для управления процессами невозможно или практически не возможно иметь дело с рациональными подгруппами. Время или стоимость, необходимые для измерения при одиночном наблюдении, настолько большие, что проведения повторных наблюдений даже не рассматривают. Это обычно происходит, когда измерение дорого стоит (например, при разрушительном контроле) или выход продукции все время относительно однороден. В других ситуациях возможно получить только одно значение, например показания прибора или значения характеристики партии исходных материалов. В таких случаях необходимо, чтобы контроль процесса базировался на индивидуальных значениях.

В случае карт индивидуальных значений для получения оценок изменчивости в пределах партии не используют рациональные подгруппы, а контрольные пределы рассчитывают на основе меры вариации, полученной по скользящим размахам, часто двух наблюдений. Скользящий размах - это абсолютное значение разницы измерений в последовательных парах; то есть это разница первого и второго измерения, потом второго и третьего и тому подобное.

Карты Шухарта

На основе скользящих размахов вычисляют средний скользящий размах R , который используют для построения контрольных карт. Также на основе всех данных вычисляют общее среднее X . В таблице приведены формулы для расчета контрольных пределов для карт индивидуальных значений.

Стати с-тика	Стандартные значения не задано		Стандартные значения задано		
	Средняя линия	ВКГ и НКГ	Средняя линия	ВКГ	НКГ
X	\bar{X}	$\bar{X} \pm E_2 \cdot \bar{R}$	$X_0 (\mu)$	$X_0 \pm 3 \cdot \sigma_0$	
R	\bar{R}	$D_3 \cdot \bar{R}$ $D_4 \cdot \bar{R}$	R_0 ($d_2 \sigma_0$)	$D_2 \sigma_0$	$D_1 \sigma_0$
$E_2 = \frac{3}{d_2}$ при n=2					

Карты Шухарта

Во время использования карт индивидуальных значений необходимо учитывать тот факт, что :

- а) карты индивидуальных значений не так чувствительны к изменениям процесса, как \bar{X} - и R –карты (карты средних арифметических);
- б) во время интерпретации карт индивидуальных значений нужно проявлять осторожность, если распределение процесса не является нормальным;
- в) Карты индивидуальных значений не выделяют повторяемость процесса от элемента к элементу, и потому в некоторых случаях лучше использовать обычные \bar{X} – и R -карты с малыми объемами выборочных подгрупп (от 2 до 4), даже если это требует увеличения интервала между подгруппами.

Карты Шухарта

A	A_2	A_3	B_3	B_4	B_5	B_6	D_1	D_2	D_3	D_4	C_4	$1/C_4$	d_2	$1/d_2$
2,1210	1,8800	2,6590	0,0000	3,2670	0,0000	2,6060	0,0000	3,6860	0,0000	3,2670	0,7979	1,2533	1,1280	0,8865
1,7320	1,0230	1,9540	0,0000	2,5680	0,0000	2,2760	0,0000	4,3580	0,0000	2,5740	0,8862	1,1284	1,6930	0,5907
1,5000	0,7290	1,6280	0,0000	2,2660	0,0000	2,0880	0,0000	4,6980	0,0000	2,2820	0,9213	1,0854	2,0590	0,4857
1,3420	0,5770	1,4270	0,0000	2,0890	0,0000	1,9640	0,0000	4,9180	0,0000	2,1140	0,9400	1,0638	2,3260	0,4299
1,2250	0,4830	1,2870	0,0300	1,9700	0,0290	1,8740	0,0000	5,0780	0,0000	2,0040	0,9515	1,0510	2,5340	0,3946
1,1340	0,4190	1,1820	0,1180	1,8820	0,1130	1,8060	0,2040	5,2040	0,0760	1,9240	0,9594	1,0423	2,7040	0,3698
1,0610	0,3730	1,0990	0,185	1,8150	0,1790	1,7510	0,3880	5,3060	0,1360	1,8640	0,9650	1,0363	2,8470	0,3512
1,0000	0,3370	1,0320	0,2390	1,7610	0,2320	1,7070	0,5470	5,3930	0,1840	1,8160	0,9693	1,0317	2,9700	0,3367
0,9490	0,3080	0,9750	0,2840	1,7160	0,2760	1,6690	0,6870	5,4690	0,2230	1,7770	0,9727	1,0281	3,0780	0,3249
0,9050	0,2850	0,9270	0,3210	1,6790	0,3130	1,6370	0,8110	5,5350	0,2560	1,7440	0,9754	1,0252	3,1730	0,3152
0,8660	0,2660	0,8860	0,3540	1,6460	0,3460	1,6100	0,9220	5,5940	0,2830	1,7170	0,9776	1,0229	3,2580	0,3069
0,8320	0,2490	0,8500	0,3820	1,6180	0,3740	1,5850	1,0250	5,6470	0,3070	1,6930	0,9794	1,0210	3,3360	0,2998
0,8020	0,2350	0,8170	0,4060	1,5940	0,3990	1,5630	1,1180	5,6960	0,3280	1,6720	0,9810	1,0194	3,4070	0,2935
0,7750	0,2230	0,7890	0,4280	1,5720	0,4210	1,5440	1,2030	5,7410	0,3470	1,6530	0,9823	1,0180	3,4720	0,2880
0,7500	0,2120	0,7630	0,4480	1,5520	0,4400	1,5260	1,2820	5,7820	0,3630	1,6370	0,9835	1,0168	3,5320	0,2831
0,7280	0,2030	0,7390	0,4660	1,5340	0,4580	1,5110	1,3560	5,8200	0,3780	1,6220	0,9845	1,0157	3,5880	0,2787
0,7070	0,1940	0,7180	0,4820	1,5180	0,4750	1,4960	1,4240	5,8560	0,3910	1,6080	0,9854	1,0148	3,6400	0,2747
0,6880	0,1870	0,6980	0,4970	1,5030	0,4900	1,4830	1,4870	5,8910	0,4030	1,5970	0,9862	1,0140	3,6890	0,2711
0,6710	0,1800	0,6800	0,5100	1,4900	0,5040	1,4700	1,5490	5,9210	0,4150	1,5850	0,9869	1,0133	3,7350	0,2677
0,6550	0,1730	0,6630	0,5230	1,4770	0,5160	1,4590	1,6050	5,9510	0,4250	1,5750	0,9876	1,0126	3,7780	0,2647
0,6400	0,1670	0,6470	0,5340	1,4660	0,5280	1,4480	1,6590	5,9790	0,4340	1,5660	0,9882	1,0119	3,8190	0,2618
0,6260	0,1620	0,6330	0,5450	1,4550	0,5390	1,4380	1,7100	6,0060	0,4430	1,5570	0,9887	1,0114	3,8580	0,2592
0,6120	0,1570	0,6190	0,5550	1,4450	0,5490	1,4290	1,7590	6,0310	0,4510	1,5480	0,9892	1,0109	3,8950	0,2567
0,6000	0,1530	0,6060	0,5650	1,4340	0,5590	1,4200	1,8060	6,0560	0,4590	1,5410	0,9896	1,0105	3,9310	0,2544

Карты Шухарта

МЕТОД УПРАВЛЕНИЯ И ИНТЕРПРЕТАЦИЯ КОНТРОЛЬНЫХ КАРТ ДЛЯ КОЛИЧЕСТВЕННЫХ ДАННЫХ

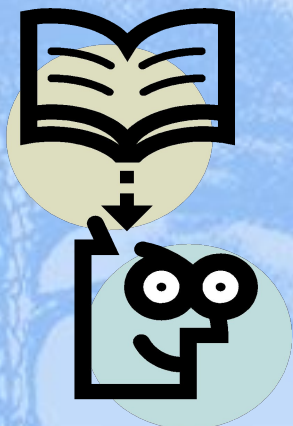
Система карт Шухарта опирается на тот факт, что когда изменчивость процесса от элемента к элементу и среднее процесса остаются постоянными на данных уровнях (оцененных, соответственно, по R и X), то размахи (R) и средние (X) отдельных подгрупп будут изменяться только случайным образом и редко выходить за контрольные пределы. Также не должно быть очевидных трендов или структур данных, кроме тех, которые возникают случайно с некоторой частью вероятности.

X -карта показывает, где находится среднее процесса и отражает стабильность процесса. X -карта выявляет нежелательные вариации между подгруппами и вариации относительно них среднего. R -карта выявляет любую нежелательную вариацию внутри подгрупп и служит индикатором изменчивости исследуемого процесса. Это мера правильности (конзистентности) и однородности процесса. Если R -карта показывает, что вариации внутри подгрупп не изменяются, то это значит, что процесс остается в статистически управляемом состоянии. Это имеет место только в том случае, если все выборки обрабатывали одинаково. Если R -карта показывает, что процесс вышел из управляемого состояния или уровень на R -карте растет, то это может обозначать, что или отдельные подгруппы испытали разную обработку, или в процессе действует несколько разных систем причинно-следственных связей.

На X -карту также могут влиять условия, при которых процесс вышел из состояния статистической управляемости на R -карте. Возможность интерпретировать размахи или средние значения подгрупп зависит от оценивания изменчивости от элемента к элементу, потому R -карту необходимо анализировать первой.

Статистика в аналитической химии

Что такое – «Неопределенность» ?



Понятие «неопределенность» (англ. «Uncertainty»), появилось более 30 лет назад. Неопределенность и связанные с ней величины (стандартная неопределенность, расширенная неопределенность и т.д.), в последнее время широко используются при представлении результатов измерений, особенно в европейских странах. То же происходит и у нас - эти термины (и соответствующие величины) применяются все чаще. Так, понятие неопределенности фигурирует в ряде нормативных документов.

Как известно, фундаментальным понятием классической теории измерений является *погрешность*

$$\delta_i = X_i - \mu$$

отклонение результата измерения X_i от истинного значения измеряемой величины μ . Погрешность возникает из-за несовершенства процесса измерений. Хотя погрешность не может быть точно известна (из - за неизвестности истинного значения), это понятие удобно использовать для

статистического описания процесса измерений. Распределение погрешности δ совпадает с распределением результатов измерений X с точностью до начала координат (см. рис 1).

Статистика в аналитической химии

Рассмотрим теперь, как определяется неопределенность. Согласно EUROCHEM/CITAC Guide “Quantifying Uncertainty in Analytical Measurements” , Second Ed., 2000, неопределенность - это «*Параметр, связанный с результатом измерения и характеризующий разброс значений, которые с достаточным основанием могут быть приписаны измеряемой величине. ... Этим параметром может быть, например, стандартное отклонение или ширина доверительного интервала*».

Согласно другим определений - неопределенность следует рассматривать как «*параметр централизованной случайной величины, представляющей собой разность между истинным значением измеряемой величины и результатом измерений, то есть величины, совпадающей по модулю с погрешностью измерений, но противоположной ей по знаку*». Другими словами, это параметр распределения величины $\mu - X_i$

Рассматривая подробнее имеющиеся литературные данные можно сказать, что **во всех случаях** в связи с неопределенностью рассматриваются **симметричные распределения результатов измерений** (результаты измерений рассматриваются как выборка из нормально распределенной генеральной совокупности), а в качестве параметра, упомянутого в определении неопределенности, **всегда** рассматриваются **стандартные отклонения**, характеризующие случайные и скрытые (невыявленные) ошибки измерений, оцененные разными способами (стандартная неопределенность, суммарная стандартная неопределенность), либо кратные им величины (расширенная неопределенность). Фактически предполагается нормальное распределение результатов измерений.

Статистика в аналитической химии

Обсуждение погрешностей данное выше известно как "классический подход" к измерению. Он хорошо служил науке об измерениях, но он основан на допущении, что измерение в конечном счете может быть описано единственным истинным значением. В последние годы произошел сдвиг в понимании того, что концепция «истинного» значения может быть не корректной, а поэтому понятия правильности, случайной и систематической погрешностей также могут быть не правомерными. «Подход с точки зрения неопределенности» принимает, то, что есть только одна неопределенность эксперимента являющаяся следствием различных компонентов. Неопределенность характеризует степень с которой значение измеряемой величины известно после измерения, принимая во внимание данную информацию от измерения.

Основные причины, по которым вводится понятие «неопределенность», следующие.

1. Отсчет доверительного интервала в отсутствие систематических погрешностей ведется от \bar{x} среднего значения результатов измерений X_i . Вообще говоря, при использовании понятия «погрешность» отсчет должен был бы вестись от математического ожидания μ^2 , а величину \bar{x} используют не от хорошей жизни – истинное значение неизвестно, а \bar{x} – есть наилучшая оценка для μ . Использование понятия «неопределенность» с этой точки зрения более логично - в определениях всех рассчитываемых параметров фигурируют только наблюдаемые величины.
2. Способы оценки интервала, в котором лежит истинная величина, более разнообразны и детально прописаны в документах, использующих понятие «неопределенность». В частности, учитываются реально имеющие место, но зачастую игнорируемые в отечественных нормативных документах скрытые, или невыявленные, систематические ошибки.

Статистика в аналитической химии

3. Использование «неопределённости» позволяет наглядно решать вопрос о соответствии (или несоответствии) измеренной характеристики качества установленным нормам. Если значение нормы не перекрывается расширенной неопределённостью результата измерения, то, основываясь на этом результате можно делать надёжное заключение о соответствии (или несоответствии) объекта испытания этой норме. Правда, то же относится к корректно рассчитанному доверительному интервалу.
4. Понятие «погрешность» исходно не являлось на Западе настолько же привычным, как у нас. Поставленные перед необходимостью оценивать интервал, в котором лежит истинное значение, зарубежные специалисты в качестве основы взяли «неопределенность» - у них был выбор.

Графические методы выяснения источников неопределенности

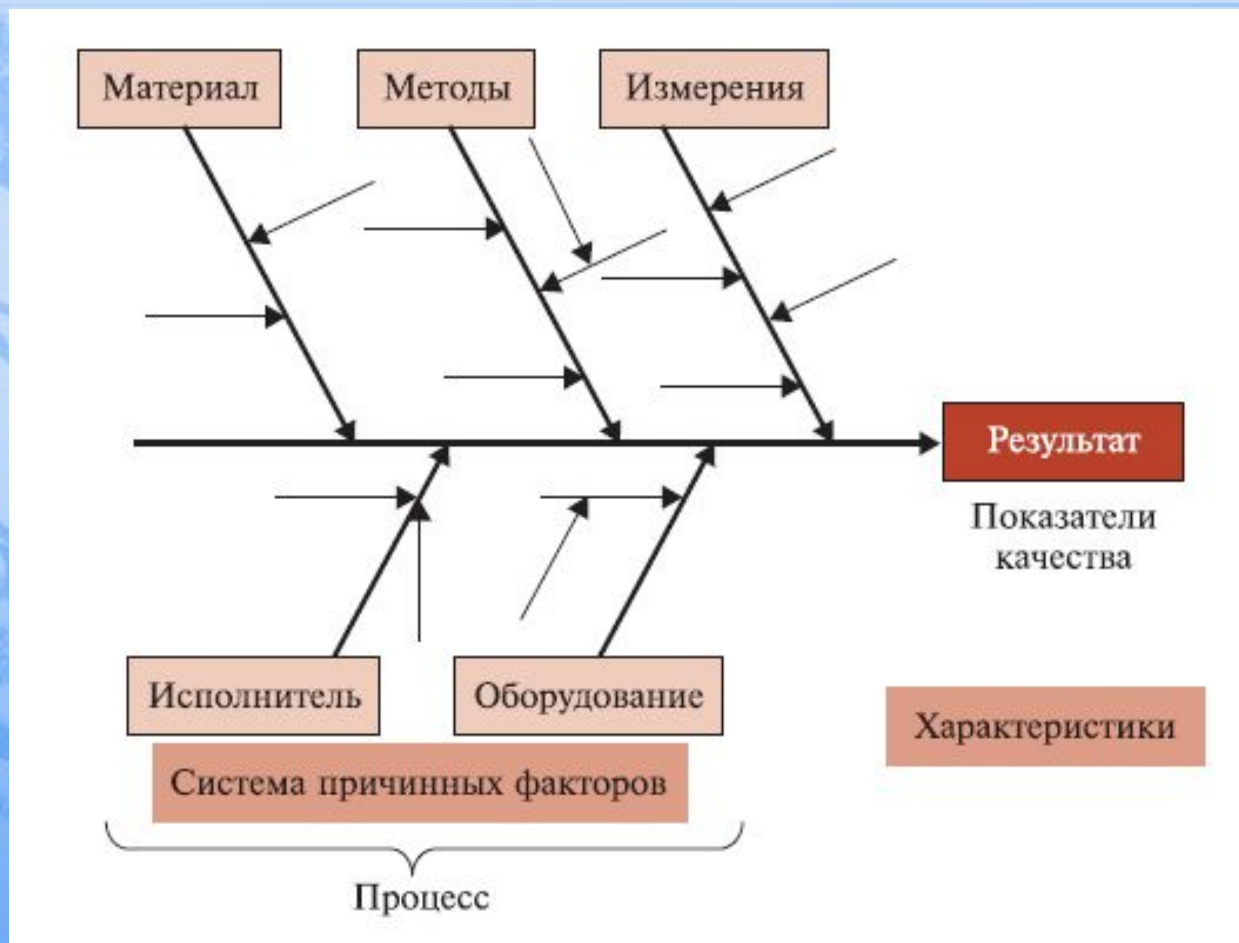


Диаграмма Исикавы

Графические методы выяснения источников неопределенности

"Причинно-следственная диаграмма" («рыбий скелет»)

Автор метода: К. Исикава (Япония), 1952 г.

Назначение метода

Применяется при разработке и непрерывном совершенствовании продукции (метода). Диаграмма Исикавы - инструмент, обеспечивающий системный подход к определению фактических причин возникновения проблем.

Цель метода

Изучить, отобразить и обеспечить технологию поиска истинных причин рассматриваемой проблемы для эффективного их разрешения.

Суть метода

Причинно-следственная диаграмма - это ключ к решению возникающих проблем. Диаграмма позволяет в простой и доступной форме систематизировать все потенциальные причины рассматриваемых проблем, выделить самые существенные и провести поуровневый поиск первопричины.

План действий

В соответствии с известным принципом Парето, среди множества потенциальных причин (причинных факторов, по Исикаве), порождающих проблемы (следствие), лишь две-три являются наиболее значимыми, их поиск и должен быть организован.

Для этого осуществляется:

- сбор и систематизация всех причин, прямо или косвенно влияющих на исследуемую проблему;
- группировка этих причин по смысловым и причинно-следственным блокам;
- ранжирование их внутри каждого блока;
- анализ получившейся картины.

ГРАФИЧЕСКИЕ МЕТОДЫ ВЫЯСНЕНИЯ ИСТОЧНИКОВ НЕОПРЕДЕЛЕННОСТИ

Общие правила построения

- Прежде чем приступать к построению диаграммы, все участники должны прийти к единому мнению относительно формулировки проблемы.
- Изучаемая проблема записывается с правой стороны в середине чистого листа бумаги и заключается в рамку, к которой слева подходит основная горизонтальная стрелка - "хребет" (диаграмму Исикавы из-за внешнего вида часто называют "рыбьим скелетом").
- Наносятся главные причины (причины уровня 1), влияющие на проблему, - "большие кости". Они заключаются в рамки и соединяются наклонными стрелками с "хребтом".
- Далее наносятся вторичные причины (причины уровня 2), которые влияют на главные причины ("большие кости"), а те, в свою очередь, являются следствием вторичных причин. Вторичные причины записываются и располагаются в виде "средних костей", примыкающих к "большим". Причины уровня 3, которые влияют на причины уровня 2, располагаются в виде "мелких костей", примыкающих к "средним", и т. д. (Если на диаграмме приведены не все причины, то одна стрелка оставляется пустой).
- При анализе должны выявляться и фиксироваться все факторы, даже те, которые кажутся незначительными, так как цель схемы - отыскать наиболее правильный путь и эффективный способ решения проблемы.
- Причины (факторы) оцениваются и ранжируются по их значимости, выделяя особо важные, которые предположительно оказывают наибольшее влияние на показатель качества.
- В диаграмму вносится вся необходимая информация: ее название; наименование изделия; имена участников; дата и т. д.

ГРАФИЧЕСКИЕ МЕТОДЫ ВЫЯСНЕНИЯ ИСТОЧНИКОВ НЕОПРЕДЕЛЕННОСТИ

Достоинства метода

Диаграмма Исикавы позволяет:

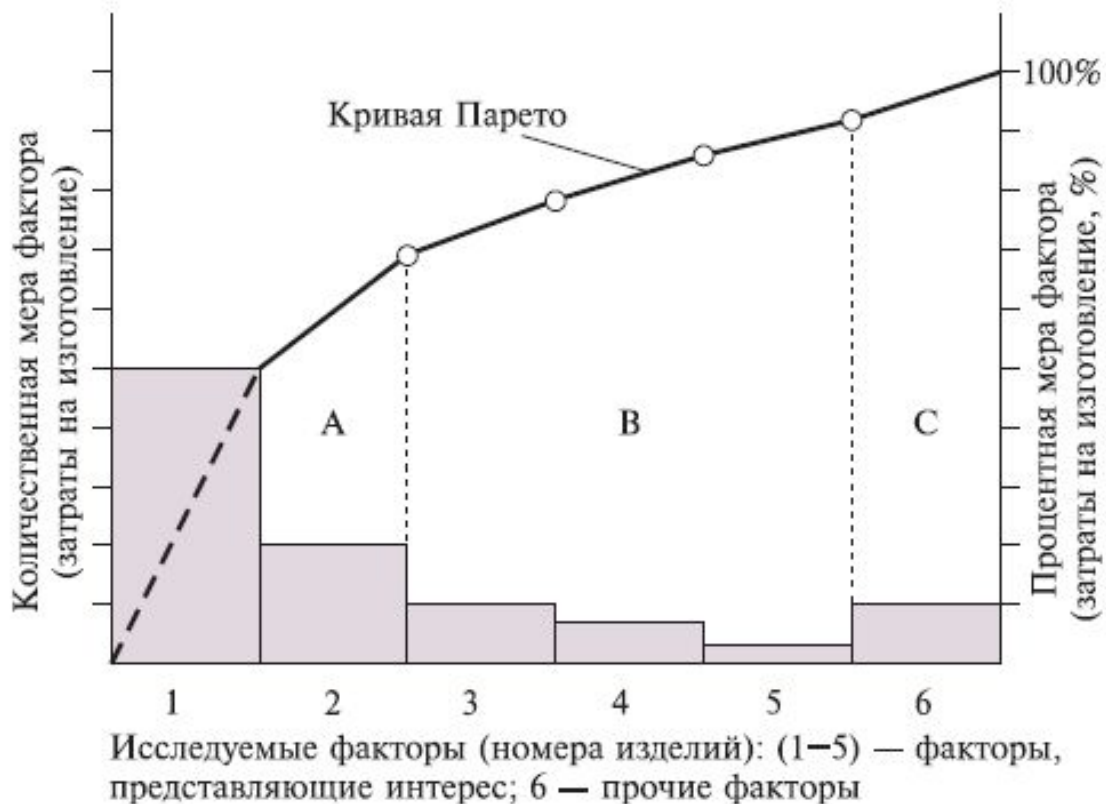
- стимулировать творческое мышление;
- представить взаимосвязь между причинами и сопоставить их относительную важность.

Недостатки метода

Не рассматривается логическая проверка цепочки причин, ведущих к первопричине, т. е. отсутствуют правила проверки в обратном направлении от первопричины к результатам.

Сложная и не всегда четко структурированная диаграмма не позволяет делать правильные выводы.

ГРАФИЧЕСКИЕ МЕТОДЫ ВЫЯСНЕНИЯ ИСТОЧНИКОВ НЕОПРЕДЕЛЕННОСТИ



ГРАФИЧЕСКИЕ МЕТОДЫ ВЫЯСНЕНИЯ ИСТОЧНИКОВ НЕОПРЕДЕЛЕННОСТИ

Авторы метода: В. Парето (Италия), 1897 г, М. Лоренц (США), 1979 г.

Назначение метода

Применяется практически в любых областях деятельности. Японский союз ученых и инженеров в 1979 г. включил диаграмму Парето в состав семи методов контроля качества.

Цель метода

Выявление проблем, подлежащих первоочередному решению.

Суть метода

Диаграмма Парето - инструмент, позволяющий выявить и отобразить проблемы, установить основные факторы, с которых нужно начинать действовать, и распределить усилия с целью эффективного разрешения этих проблем.

Различают два вида диаграмм Парето:

- по результатам деятельности - предназначена для выявления главной проблемы нежелательных результатов деятельности;
- по причинам - используется для выявления главной причины проблем, возникающих в ходе производства.

План действий

- Определить проблему, которую надлежит решить.
- Учесть все факторы (признаки), относящиеся к исследуемой проблеме.
- Выявить первопричины, которые создают наибольшие трудности, собрать по ним данные и проранжировать их.
- Построить диаграмму Парето, которая объективно представит фактическое положение дел в понятной и наглядной форме.
- Провести анализ диаграммы Парето.

ГРАФИЧЕСКИЕ МЕТОДЫ ВЫЯСНЕНИЯ ИСТОЧНИКОВ НЕОПРЕДЕЛЕННОСТИ

Общие правила построения диаграммы Парето

- Решить, какие проблемы (причины проблем) надлежит исследовать, какие данные собирать и как их классифицировать.
- Разработать формы для регистрации исходных данных (например, контрольный листок).
- Собрать данные, заполнив формы, и подсчитать итоги по каждому исследуемому фактору (показателю, признаку).
- Для построения диаграммы Парето подготовить бланк таблицы, предусмотрев в нем графы для итогов по каждому проверяемому фактору в отдельности, накопленной суммы числа появлений соответствующего фактора, процентов к общему итогу и накопленных процентов.
- Заполнить таблицу, расположив данные, полученные по проверяемому фактору, в порядке убывания значимости.
- Подготовить оси (одну горизонтальную и две вертикальные линии) для построения диаграммы. Нанести на левую ось ординат шкалу с интервалами от 0 до общей суммы числа выявленных факторов, а на правую ось ординат - шкалу с интервалами от 0 до 100, отражающую процентную меру фактора. Разделить ось абсцисс на интервалы в соответствии с числом исследуемых факторов или относительной частотой.
- Построить столбиковую диаграмму. Высота столбца (откладывается по левой шкале) равна числу появлений соответствующего фактора. Столбцы располагают в порядке убывания (уменьшения значимости фактора). Последний столбец характеризует "прочие", т. е. малозначимые факторы, и может быть выше соседних.
- Начертить кумулятивную кривую (кривую Парето) - ломаную, соединяющую точки накопленных сумм (количественной меры факторов или процентов). Каждую точку ставят над соответствующим столбцом столбиковой диаграммы, ориентируясь на его правую сторону.
- Нанести на диаграмму все обозначения и надписи.
- Провести анализ диаграммы Парето.

ГРАФИЧЕСКИЕ МЕТОДЫ ВЫЯСНЕНИЯ ИСТОЧНИКОВ НЕОПРЕДЕЛЕННОСТИ

Достоинства метода

- Простота и наглядность делают возможным использование диаграммы Парето специалистами, не имеющими особой подготовки.
- Сравнение диаграмм Парето, описывающих ситуацию до и после проведения улучшающих мероприятий, позволяют получить количественную оценку выигрыша от этих мероприятий.

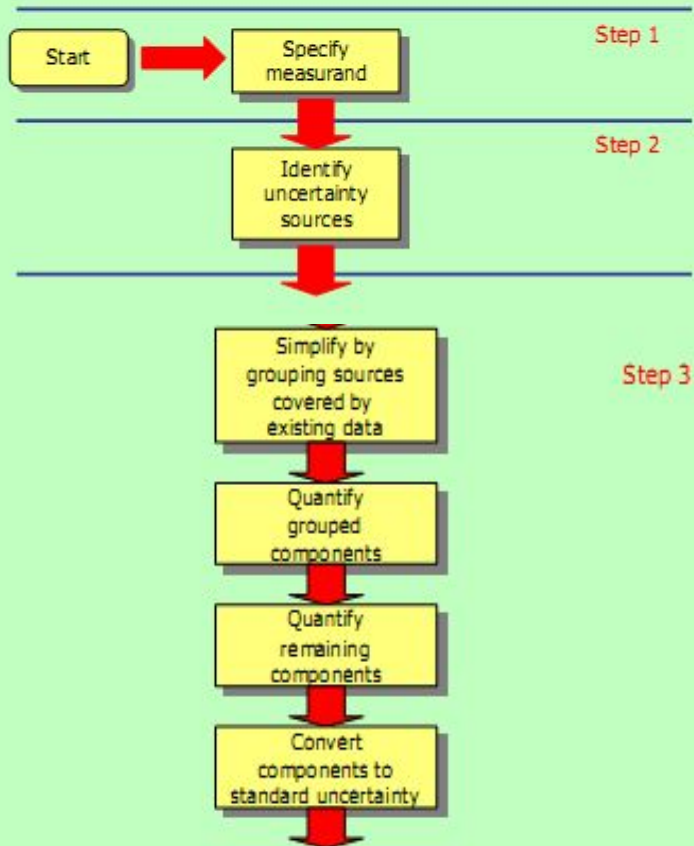
Недостатки метода

При построении сложной, не всегда четко структурированной диаграммы возможны неправильные выводы.

ПРОЦЕСС ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ СОГЛАСНО РУКОВОДСТВА EURACHEM /CITAC

The Uncertainty Estimation Process

- The Uncertainty Estimation Process as described in the EURACHEM/CITAC-guide can be divided in 4 steps which are shown graphically below



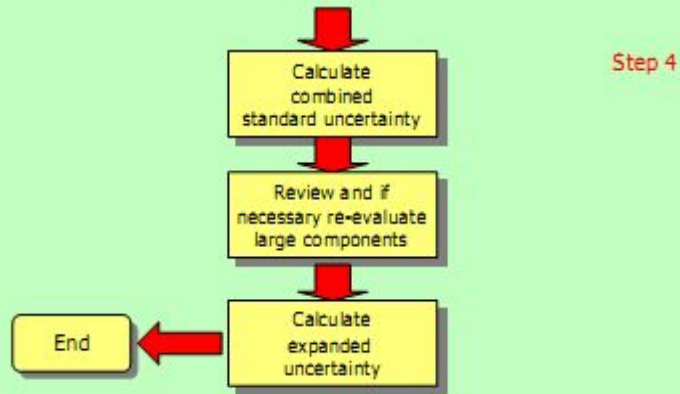
Процесс оценивания неопределенности

- Процесс оценивания неопределенности описан в руководстве EURACHEM/CITAC и состоит из четырех основных последовательных шагов.



ПРОЦЕСС ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ СОГЛАСНО РУКОВОДСТВА EURACHEM /CITAC

The Uncertainty Estimation Process



Шаг 4

Расчитать суммарную стандартную неопределенность

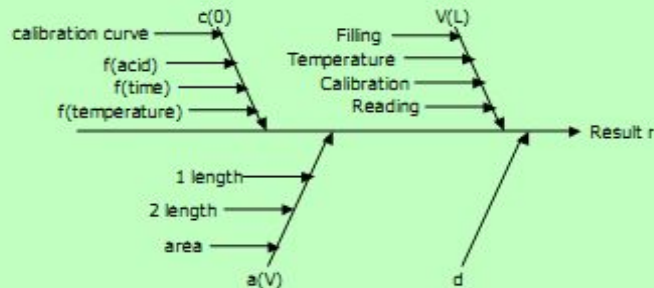
При необходимости пересчитать крупные компоненты

Расчитать расширенную неопределенность

Конец

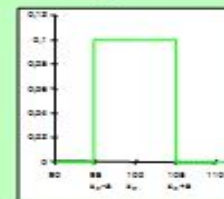
The Uncertainty Estimation Process Step 2

- Identification of uncertainty sources
 - Could be described e.g. by a fishbone diagram

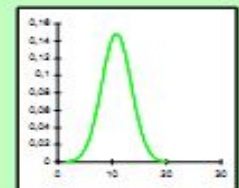


The Uncertainty Estimation Process Step 3

- Quantification of uncertainties
 - Try to find an uncertainty for each component
 - Some are normal distributed
 - Others may have a rectangular distribution



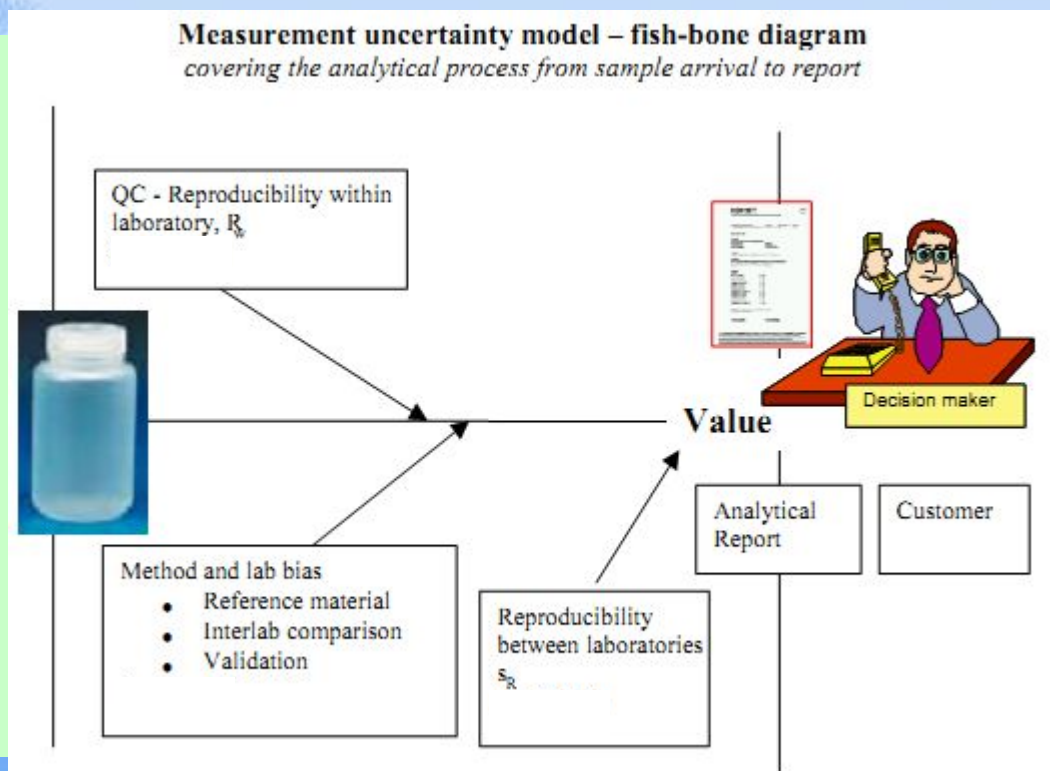
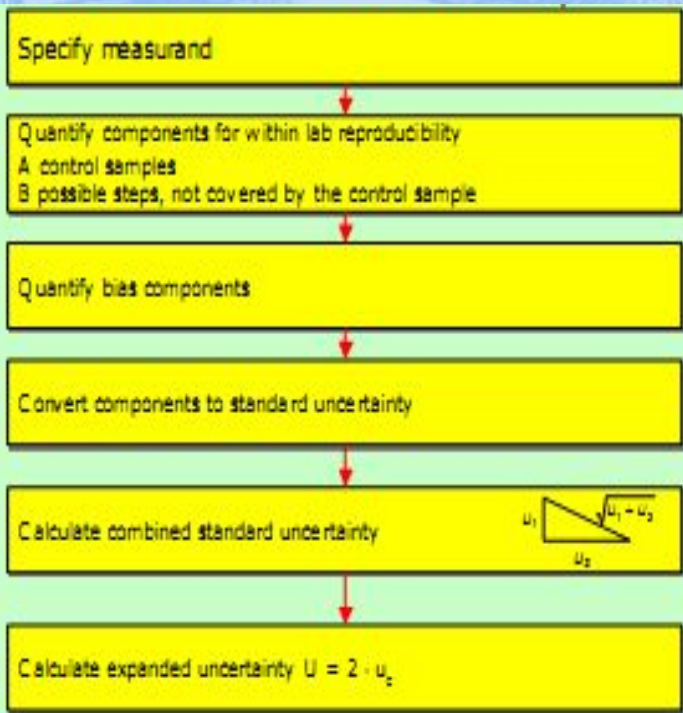
$$s = \sqrt{\frac{a^2}{3}}$$



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

АЛЬТЕРНАТИВНЫЙ ПРОЦЕСС ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ СОГЛАСНО РУКОВОДСТВА **NORDTEST TR 537**

В руководстве **NORDTEST TR 537** «Handbook for calculation of measurement uncertainty in Environmental laboratories» приводится альтернативный процесс оценки неопределенности, базирующийся на данных, которые учитывают результаты межлабораторных испытаний и валидации методов (внутрилабораторная воспроизводимость, робасность).



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ



Рекомендации ISO/IEC 17025:2005 по оцениванию неопределенности измерений в испытательных и калибровочных лабораториях

Стандарт ISO/IEC 17025:2006 определяет международное признание результатов испытаний и калибровок лабораториями, получивших аккредитацию от органов, которые заключили MRA¹ с аналогичными органами других стран. Он законодательно закрепил необходимость наличия процедур оценивания неопределенности измерений, проводимых в аккредитованных лабораториях

- а)** при выборе, разработке и оценивании пригодности методов и процедур, которые используются в деятельности лаборатории (п. 5.4.1);
- б)** при использовании стандартизованных и нестандартизованных или разработанных лабораторией методов и процедур калибровки или испытания (п. 5.4.6);
- в)** при оформлении свидетельств о калибровке и протоколов испытаний (п.5.6.2.1.1. 5.10.4.1);
- г)** при создании программ и процедур калибровки своих собственных исходных эталонов, образцовых веществ и оборудования, для обеспечения прослеживаемости проводимых лабораторией калибровок и измерений до Международной системы единиц (SI) (п. 5.6).

MRA - The Mutual Recognition Arrangement - соглашение о взаимном признании национальных эталонов и сертификатов о калибровке и измерениях, выпускаемых национальными метрологическими институтами, подготовленное Международным комитетом по мерам и весам (МКМВ)



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Основные термины и положения по оцениванию неопределенности измерений, согласно *GUM «Guide to the Expression of Uncertainty in Measurement» ISO, Geneva, First Edition. 1995* и *EURACHEM / CITAC Guide Quantifying Uncertainty Analytical Measurement, Second Ed., 2002.*

Неопределенностью измерения называется параметр, связанный с результатом измерений и характеризующий рассеяние значений, которые могут быть обоснованно приписаны измеряемой величине.

Приведенное выше определение термина *неопределенность* сосредотачивает внимание на интервале значений, которые, как полагает аналитик, могут быть обоснованно приписаны измеряемой величине.

Источники неопределенности

На практике неопределенность результата измерения может возникать вследствие влияния многих возможных источников, включая, например, такие как неполное определение измеряемой величины, пробоотбор, эффекты матрицы и мешающие влияния, условия окружающей среды, погрешности средств измерений массы и объема, неопределенности значений эталонов, приближения и допущения, являющиеся частью метода и процедуры измерений, а также случайные колебания.



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Стандартная неопределенность — неопределенность результата измерений, выраженная как стандартное отклонение.

Суммарная стандартная неопределенность - стандартная неопределенность результата измерений, когда результат получают из значений ряда других величин, равная положительному квадратному корню суммы членов, причем члены являются дисперсиями или ковариациями этих других величин, взвешенными в соответствии с тем, как результат измерений изменяется в зависимости от изменения этих величин.

Оценка (неопределенности) по типу A - метод оценивания неопределенности путем статистического анализа ряда наблюдений.

Оценка (неопределенности) по типу B - метод оценивания неопределенности иным способом, чем статистический анализ ряда наблюдений.

Расширенная неопределенность - величина, определяющая интервал вокруг результата измерений, в пределах которого, можно ожидать, находится большая часть распределения значений, которые с достаточным основанием могли бы быть приписаны измеряемой величине.

Коэффициент охвата - числовой коэффициент, используемый как множитель суммарной стандартной неопределенности для получения расширенной неопределенности.



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Основные положения концепции неопределенности :

1. Все составляющие неопределенности в результате измерения можно сгруппировать в две категории **в соответствии со способом их оценивания**:

A - составляющие, оцениваемые путем применения статистических методов (обработкой результатов многократных измерений);

B - составляющие, оцениваемые другим способом (по характеристикам, взятым из предыдущих экспериментов, из паспорта на прибор, методик выполнения измерений, из справочников и т.д.). К источникам этих неопределенностей относятся: случайные погрешности однократных измерений; поправки, вводимые в результат измерения, неисключенные систематические погрешности, погрешности справочных данных и т.д.

2. Составляющие типа **A** оцениваются как стандартные (среднеквадратические) отклонения средних арифметических многократных наблюдений (неопределенности типа **A** - u_A). Эти составляющие характеризуются числами степеней свободы ν (для прямых многократных измерений $\nu = n - 1$, где n - количество многократных наблюдений, выполненных при измерении). Если между составляющими типа **A** наблюдается статистическая взаимосвязь (корреляция), то необходимо указывать коэффициенты корреляции $r(x_i, x_j) = r_{ij}$.

3. Составляющие типа **B** также оцениваются как **предполагаемые** стандартные (среднеквадратические) отклонения, получаемые из известных границ, в которых могут находиться значения этих составляющих (неопределенности типа **B** - u_B). Между составляющими типа **B** может существовать **предполагаемая** взаимосвязь, тогда следует указывать коэффициенты предполагаемой корреляции.

4. Суммарная стандартная (среднеквадратическая) неопределенность u_c рассчитывается по правилу суммирования дисперсий :

СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

$$u_c = \sqrt{u_A^2 + u_B^2}$$

5. Интервальной оценкой неопределенности является расширенная неопределенность U , которую получают путем умножения стандартной суммарной неопределенности u_c на так называемый коэффициент охвата k . При указании расширенной неопределенности указывают уровень доверия вероятность P . Обычно эта вероятность принимается равной 0.95.

Основные расчетные формулы для определения неопределенности

1. Стандартная неопределенности типа **A**

$$u_A(x_i) = \sqrt{\frac{\sum_{q=1}^{n_i} (x_{iq} - \bar{x}_i)^2}{n_i \cdot (n_i - 1)}}$$

2. Стандартная неопределенность типа **B**

$$u_B(x_i) = \frac{\theta_i}{\sqrt{\alpha_i}}$$

где, θ_i – неисключенная систематическая погрешность, α_i – коэффициент, связанный принимаемым законом распределения ($\alpha_i = 3$ для равномерного, $\alpha_i = 6$ для треугольного, $\alpha_i = 2$ для арксинуса и $\alpha_i = 4$ для нормального закона распределения)



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

3. Коэффициенты чувствительности. Расчет их необходим для определения вклада неопределенности каждой входной величины в неопределенность измеряемой

$$c_i = \frac{\partial f}{\partial x_i} = \frac{\partial Y}{\partial X_i} \Big|_{x_1, x_2, \dots, x_m}$$

4. Вклад неопределенности каждой входной величины в неопределенность измеряемой величины.

$$u_i(y) = c_i \cdot u(x_i)$$

5. Суммарная стандартная неопределенность

$$u(y) = \sqrt{\sum_{i=1}^m c_i^2 \cdot u^2(x_i)}$$

6. Эффективное число степеней свободы

$$v_{eff} = \frac{u^4(y)}{\sum_{i=1}^m u_i^4(y) / v_i}$$

7. Коэффициент охвата

$$k = t_{(0.05, v_{eff})}$$

8. Расширенная неопределенность

$$U = k \cdot u(y)$$



Современные методы оценки неопределенности

Базовый алгоритм оценивания неопределенности в измерениях



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Важным фактором корректной оценки неопределенности является составление модельного уравнения. Модельным уравнением является зависимость между выходной и входными величинами. Как входные величины, в модельное уравнение входят как измеренные величины, так и разные поправки и коэффициенты. Расширенную и суммарную стандартную неопределенность выходной величины образует неопределенность входных величин.

Значения входных величин находят путем измерения с однократным или многократным наблюдением, или оцениванием, из внешних источников (паспортные данные на мерную посуду, весы и тому подобное). Определяют типы оценивания неопределенности (А или В) и законы распределения входных величин.

Составление бюджета неопределенности

Вычисленные вклады неопределенности удобно представлять в виде бюджета неопределенности, который включает в себя список всех входных величин X_1, \dots, X_m их оценок x_1, \dots, x_m вместе с принадлежащими им стандартными неопределенностями измерения $u(x_i)$ и законами их распределения, а также числами степеней свободы. Кроме этого для каждой величины таблица должна содержать вклад неопределенности.

Для занесенных в таблицу числовых значений должны указываться единицы измерения для соответствующей величины. В нижней строке бюджета неопределенности можно расположить информацию о выходной величине (входная величина Y , ее оценка y , неопределенность входной величины $u(y)$, эффективное число степеней свободы ν_{eff} , коэффициент охвата k , расширенная неопределенность U).

Бюджет неопределенности представляет собой таблицу и является удобной основой для составления программы для оценивания неопределенности измерений в среде Excel.

СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

	A	B	C	D	E	F	G
12	Исходные данные Y	Исходные данные X (расчет по калибровке)		Калибровка исходные данные по Y	Калибровка исходные данные по X	Sy	Коэффициенты регрессии
13	11841.17	4.0539066		6728.29	1.013777525	32.682	a, [y.e]
14	11815.93	4.0388989		8337.99	1.970908498		5023.322
15	11435.39	3.8126290		10041.33	2.983718011		
16	11703.49	3.9720418		11829.33	4.046866536		
17	11854.5	4.0618327		13433.51	5.000715306		
18	11503.67	3.8532284		15086.585	5.983637145		
19	11908.67	4.0940423		Модельное уравнение измерения $\omega = \frac{x \cdot V_0}{V_{al} \cdot m \cdot 10} \quad x = \frac{(S - a)}{b}$			
20	11547.67	3.8793909					
21	11926.32	4.1045370					
22	среднее знач.						
23	11726.31222	3.9856					
24	СКО						
25	62.08779936	0.1108					
27	Количество измерений	9					
28	Критерий Горвица	RSDr					
29	2.088503221	2.77881053					

=(A17-\$G\$14)/\$G\$17

=СРЗНАЧ(B13:B21)

=СТАНДОТКЛОН(B13:B21)

Современные методы оценки неопределенности

	A	B	C	D	E	F	G
1	РАСЧЕТ НЕОПРЕДЕЛЕННОСТИ ИЗМЕРЕНИЯ СОДЕРЖАНИЯ Д.В. Трибенурон-метила (GUM)						
2	Входные	Значение	Границы входной	Стандартная	Число	Закон	Вклад
3	величины	входной	величины	неопределенность	степеней	распределения	неопределенности
4				входной	свободы		
5	X, [мкг]	3.9856120	0.110752605	0.036917535	8	A, нормальный	0.6922
6	V ₀ , [мл]	7.5	0.1	0.057735027	10000	B, равномерный	0.5753
7	t, [г]	0.01	0.0001	5.7735E-05	10000	B, равномерный	-0.4315
8	V _{al} , [мкл]	4	0.2	0.115470054	10000	B, равномерный	-2.1577
9	Выходная	Значение	Суммарная	Эффективное	Коэффициент	Расширенная неопределенность	
	величина	выходной	стандартная	число степеней	охвата		
		величины, %	неопределенность	свободы			
10	C, [%]	74.7302	2.3774	1034	1.9623	4.6651	

=B23

=B25

=C5/КОРЕНЬ(9)

=C6/КОРЕНЬ(3)

=(B\$5+D5/2)*B\$6/(B\$7*B\$8*10)-(B\$5-D5/2)*B\$6/(B\$7*B\$8*10)

=КОРЕНЬ(G5^2+G6^2+G7^2+G8^2)

=C10*E10

=СТЫЮДРАСПОБР(0.05;D10)

=ЦЕЛОЕ((C10^4)/(G5^4/E5+G6^4/E6+G7^4/E7+G8^4/E8))

СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

НЕДОСТАТКИ ПОДХОДА GUM К ОЦЕНКЕ НЕОПРЕДЕЛЕННОСТИ

- В основе общего подхода лежит т.н. закон распространения неопределенности, базирующийся на разложении нелинейной модельной функции в ряд Тейлора первого порядка. Применение такого подхода при существенно нелинейной зависимости дает смещенную оценку результата измерений и недостоверную оценку суммарной стандартной неопределенности.
- При нахождении расширенной неопределенности предполагается, что закон распределения выходной величины в соответствии с центральной предельной теоремой - нормальный и коэффициент охвата хорошо аппроксимируется коэффициентом Стьюдента с числом степеней свободы, определяемым формулой Велча-Саттерсвейта. Первое предположение справедливо для случая линейного модельного уравнения, большого числа входных величин и их симметричных законов распределения.
- Формула Велча-Саттерсвейта принципиально не предназначена для работы с коррелированными входными величинами и ее применение в этом случае может привести к существенной недостоверности оценивания расширенной неопределенности.

$$v_{eff} = \frac{u^4(y)}{\sum_{i=1}^m u_i^4(y)/v_i}$$



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Основываясь на указанных недостатках, было предложено применить метод статистического моделирования (метод Монте-Карло) для определения расширенной неопределенности в измерениях.

Порядок оценивания расширенной неопределенности согласно и с использованием метода статистического моделирования Монте-Карло и применением программного обеспечения MS Excel с пакетом анализа является наиболее удобным и легко осуществленным.

Он заключается в следующих операциях :

1 Формирование массивов данных

а) Генерируют L массивов (по количеству входных величин; для абсолютного градуирования $L = 7$) случайных чисел заданного объема n ($n = 10^5$), что подчиняются необходимым законам распределения. Для этого используют программу MS Excel с пакетом анализа и опцию «Генерация случайных чисел». Для входных величин, оцениваемых по типу «А», генерируют нормально распределенные массивы с заданием среднеквадратического отклонения в соответствии с оценкой каждой входной величины. Для входных величин, оцениваемых по типу «В», генерируют равномерно распределенные массивы с заданием симметричного диапазона оценок этих входных величин. То есть, если для дозирования объема V_0 используется мерная пипетка вместимостью 10 см^3 , то оценка V_0 будет составлять $\pm 0,1 \text{ см}^3$. (от $- 0,1 \text{ см}^3$ к $+ 0,1 \text{ см}^3$).

б) Формируют массив выходной величины, путем подстановки значений входных величин, которые сгенерировали, в модельное уравнение.

в) Упорядочивают (ранжируют) массив выходной величины.

СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

2. Расчет параметров бюджета неопределенности (расширенной и стандартной суммарной неопределенности, коэффициента охвата) по методу Монте-Карло

По получении массива данных выходной величины определяют:

- оценку среднего результата измерения по формуле :

$$\bar{y} = \frac{1}{n} \sum_{q=1}^n y_q$$

, где y_q - значение выходной величины;

n - объем массива выходной величины.

- оценку суммарной стандартной неопределенности результата измерения :

$$\hat{u}_c(y) = \sqrt{\frac{\sum_{q=1}^n [y_q - \bar{y}]^2}{n-1}}$$

, где n - объем массива выходной величины.

y_q - значение выходной величины;

\hat{y} - оценка среднего результата измерения.



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

- расширенную неопределенность для заданного уровня доверия. Для получения оценки расширенной неопределенности необходимо рассчитать интерквантильный интервал по формуле:

$$\hat{U}_p = \frac{1}{2} [Y_{n(1+p)/2} - Y_{n(1-p)/2}]$$

, где $Y_{n(1+p)/2}$ и $Y_{n(1-p)/2}$ -соответственно $n(1+p)/2$ и $n(1-p)/2$ члены упорядоченного массива данных выходной величины.

Для $p = 0,95$ и $n = 10^5$ квантилей распределения выходной величины оцениваются для 97500 и 2500 члена упорядоченного массива выходной величины.

- оценку коэффициента охвата по формуле :

$$\hat{k} = \hat{U}_p / \hat{u}_c(y)$$

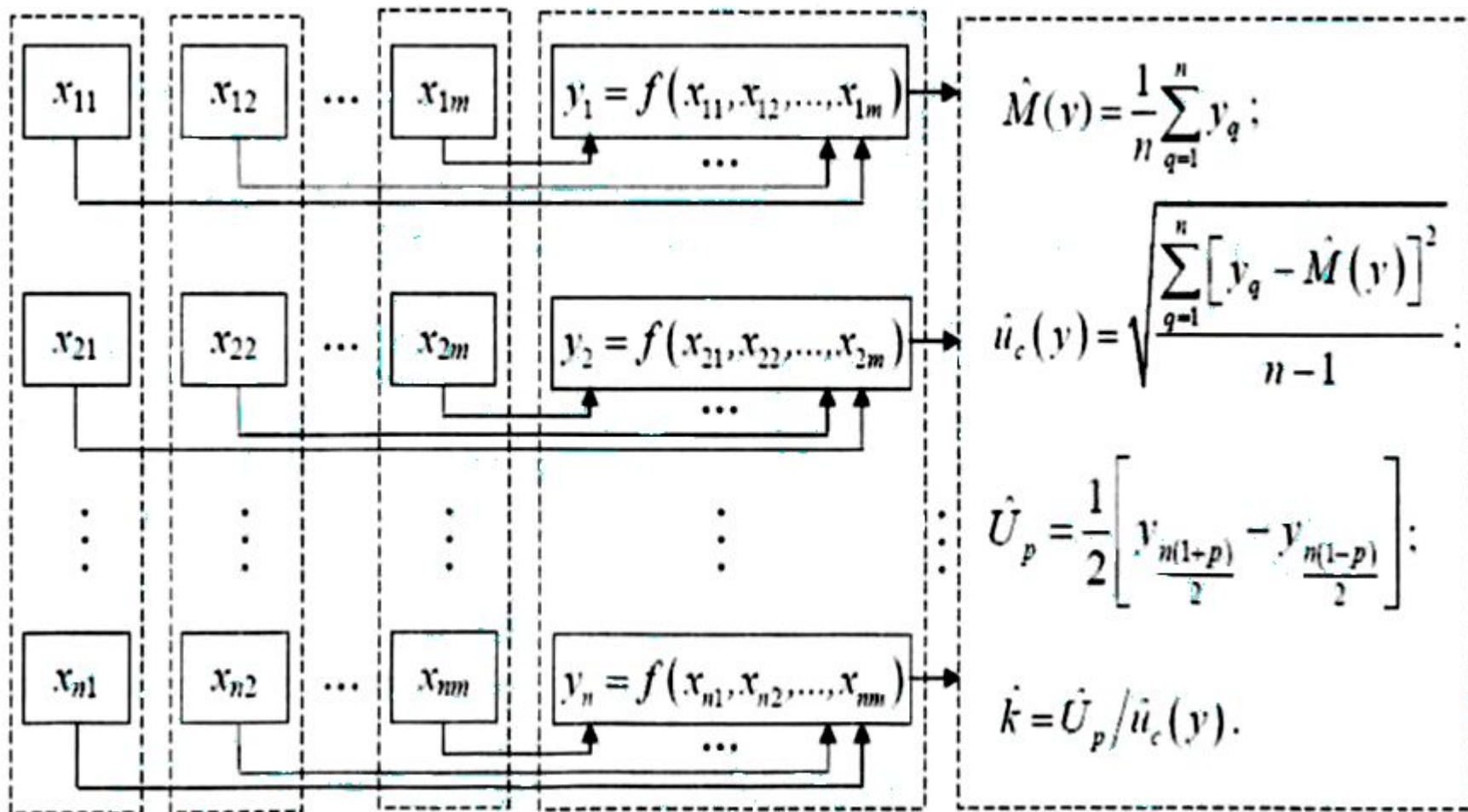
, где \hat{U}_p - расширенная неопределенность для заданного уровня доверия;

$\hat{u}_c(y)$ - оценка суммарной стандартной неопределенности.



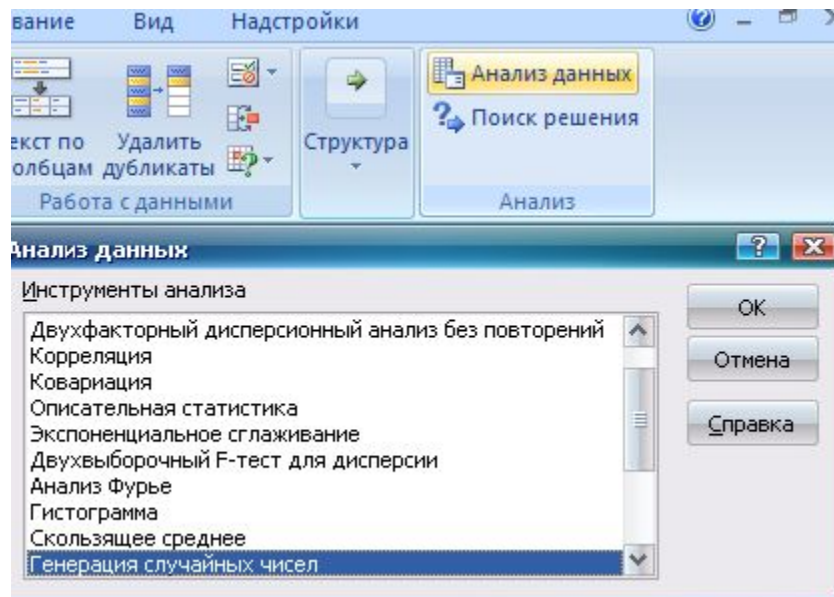
СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

СХЕМА РЕАЛИЗАЦИИ АЛГОРИТМА МОНТЕ-КАРЛО



Современные методы оценки неопределенности

При реализации метода Монте-Карло в среде Exсell устанавливают надстройку «Анализ данных», в которой вызывают функцию «Генерация случайных чисел»



Современные методы оценки неопределенности

1. Генерируют массив из 10000 данных входной величины, оцениваемой по типу **A**. В данном примере это X - масса вещества, нанесенного в точку ТСХ пластинки, получаемая в результате расчета с использованием уравнения регрессии, описывающего калибровочную зависимость между площадью хроматографического пятна и соответствующей массы вещества.

	A	B	C	D
1	<i>Расчет расширенной неопределенности</i>			
2	<i>Трибенурон-метила мет</i>			
4	Входные величины,	Входные величины, оцениваемые по типу B		
6	X , [мкг]	V_0 , [мл]	m , [г]	V_{al} , [мкл]
8	4.0299			
9	3.7540			
10	3.9980			
11	3.8956			
12	4.0549			
13	3.9497			
14	4.3057			
15	3.9634			
16	4.2699			
17	3.7899			

Генерация случайных чисел

Число переменных: 1

Число случайных чисел: 10000

Распределение: Нормальное

Параметры

Среднее = 3.9856

Стандартное отклонение = 0.03692

Случайное рассеивание:

Параметры вывода

Выходной интервал: \$A\$8

Новый рабочий лист:

Новая рабочая книга



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

2. Генерируют массив из 10000 данных входной величины, оцениваемой по типу **B**. В данном примере это V_0 – исходный объем экстракта.

	A	B	C	D
1	Расчет расширенной неопределенности			
2	Трибенурон-метила мет			
3				
4	Входные величины,	Входные величины, оцениваемые по типу B		
5				
6	X, [МКГ]	V₀,[МЛ]	m,[Г]	V_{аб},[МКЛ]
7				
8	4.0299	7.4728		
9	3.7540	7.5808		
10	3.9980	7.5784		
11	3.8956	7.5081		
12	4.0549	7.5883		
13	3.9497	7.5302		
14	4.3057	7.4869		
15	3.9634	7.5804		
16	4.2699	7.5749		
17	3.7899	7.5678		
18	3.9044	7.5565		
19	3.9314	7.4982		

Генерация случайных чисел

Число переменных: 1

Число случайных чисел: 10000

Распределение: Равномерное

Параметры

Между 7.4 и 7.6

Случайное рассеивание:

Параметры вывода

Выходной интервал: \$B\$8

Новый рабочий лист:

Новая рабочая книга

OK Отмена Справка



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

3. Продолжают генерирование массивов из 10000 данных входных величин, оцениваемых по типу **B**. В данном примере это m - масса навески препарата и V_{al} – объем экстракта, нанесенный микрошприцом в точку на ТСХ пластинке .

	A	B	C	D
1	<i>Расчет расширенной неопределенности измерения</i>			
2	<i>метила методом Монте</i>			
3				
4	Входные величины, оцениваемые по типу A		Входные величины, оцениваемые по типу B	
5				
6	X, [мкг]	V_0 , [мл]	m , [г]	V_{al} , [мкл]
8	4.0299	7.4728	0.0099	
9	3.7540	7.5808	0.0099	
10	3.9980	7.5784	0.0100	
11	3.8956	7.5081	0.0100	
12	4.0549	7.5883	0.0100	
13	3.9497	7.5302	0.0100	
14	4.3057	7.4869	0.0101	
15	3.9634	7.5804	0.0100	
16	4.2699	7.5749	0.0099	
17	3.7899	7.5678	0.0100	

Генерация случайных чисел

Число переменных: 1 ОК

Число случайных чисел: 10000 Отмена


Распределение: Равномерное Справка

Параметры

Между 0.0099 и 0.0101

Случайное рассеивание:

Параметры вывода

Выходной интервал: 

Новый рабочий лист:

Новая рабочая книга

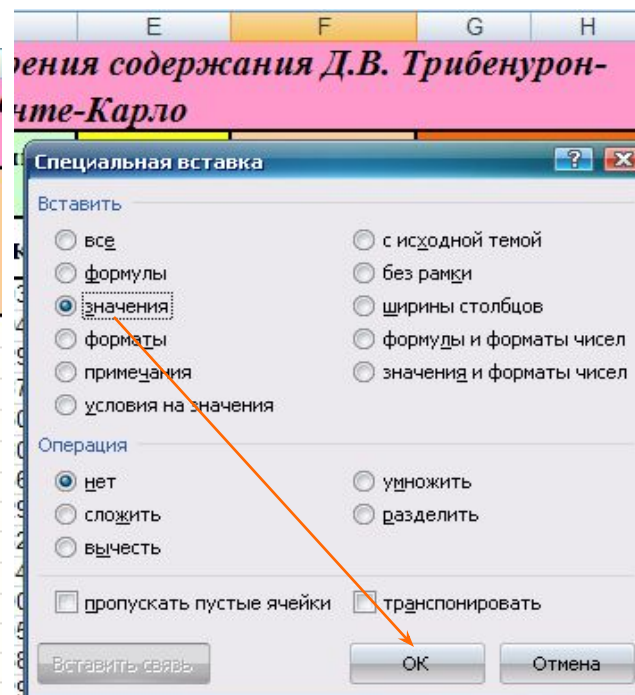


СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

4. После того как будут сгенерированы все массивы входных величин, в колонке «Е» будет выполнен расчет выходной величины (массовой доли трибенурон-метила) в препарате защиты растений, согласно модельного уравнения для всех 10000 значений. Полученные данные необходимо скопировать диапазоном «E8:E10008» и вставить в колонку «F» используя специальную вставку только «Значений», которая исключает копирование формул расчета.

$$=(A8*B8)/(C8*D8*10)$$

	A	B	C	D	E
1	Расчет расширенной неопределенности измерения содержания				
2	метила методом Монте-Карло				
4	Входные величины, оцениваемые по типу A	Входные величины, оцениваемые по типу B			Выходная величина, %
6	X, [МКГ]	V₀,[МЛ]	m, [Г]	V_{аб}, [МКЛ]	
8	4.0299	7.4728	0.0099	3.8803	78.3026
9	3.7540	7.5808	0.0099	4.0494	70.6642
10	3.9980	7.5784	0.0100	3.8229	79.3881
11	3.8956	7.5081	0.0100	3.8137	76.8804
12	4.0549	7.5883	0.0100	4.0460	75.9921
13	3.9497	7.5302	0.0100	4.1180	72.1867
14	4.3057	7.4869	0.0101	4.0456	79.1422
15	3.9634	7.5804	0.0100	3.8829	77.4114
16	4.2699	7.5749	0.0099	3.9582	82.4573
17	3.7899	7.5678	0.0100	3.8254	74.6034
18	3.9044	7.5565	0.0100	3.9790	74.3519
19	3.9314	7.4982	0.0101	3.9505	73.8948
20	4.0418	7.4205	0.0100	3.8468	78.1569
21	3.9482	7.4986	0.0101	4.0029	73.4333



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

5. Для получения результата по расчету расширенной неопределенности, скопированные данные в колонке "F" необходимо ранжировать (упорядочить) по возрастанию, используя возможности Excel. После ранжирования в ячейке «G8» будет рассчитан интерквантильный диапазон, обеспечивающий получение значения расширенной неопределенности. Суммарная стандартная неопределенность рассчитывается как СКО из 10000 значений выходной величины. Коэффициент охвата находят путем деления значения расширенной неопределенности на величину суммарной стандартной неопределенности.

$$=G8/E1000$$

8

$$=(F9758-F258)/$$

2

	A	B	C	D	E	F	G	H
1	Расчет расширенной неопределенности измерения содержания Д.В. Трибенурон-							
2	метила методом Монте-Карло							
4	Входные величины, оцениваемые по типу А		Входные величины, оцениваемые по типу В		Выходная величина, %	Упорядоченная выходная величина	Расширенная неопределенность	
6	X, [МКГ]	V₀,[МЛ]	m,[Г]	V_{al},[МКЛ]				
8	4.0004	7.4728	0.0099	3.8803	77.7285	66.0563		4.328388796
9	3.9084	7.5808	0.0099	4.0494	73.5706	68.1710		
10	3.9897	7.5784	0.0100	3.8229	79.2246	68.5861		Коэффициент охвата
11	3.9556	7.5081	0.0100	3.8137	78.0650	68.6950		1.815735849
12	4.0087	7.5883	0.0100	4.0460	75.1261	68.8215		
13	3.9736	7.5302	0.0100	4.1180	72.6238	68.9801		Суммарная стандартная неопределенность
14	4.0923	7.4869	0.0101	4.0456	75.2195	69.0403		
15	3.9782	7.5804	0.0100	3.8829	77.6999	69.1229		
16	4.0804	7.5749	0.0099	3.9582	78.7972	69.1444		2.383820751

$$=E1000$$

8

Принцип расчета неопределенности согласно NORDTEST TR 537

Шаг	Действие	Образец аммонийный азот NH4-N
1	Определить параметры	Аммонийный азот в воде определяют соответственно EN/ISO 11732/11. Заказчик заявляет о уровне неопределенности $\pm 10\%$
2	Определить значение R_w А) через контрольный образец; Б) через шаги неохваченные контрольным образцом	А) Контрольный предел отклонения установлен $\pm 3,34\%$ Б) Контрольный образец включает все аналитические шаги
3	Определить значение смещения при выполнении анализа	Это значение находят из результатов межлабораторных испытаний для лаборатории, усредняя их и из значения межлабораторной воспроизводимости. Например, межлабораторные сравнений за последние 3 года результаты смещения были 2.4; 2.7; 1.9; 1.4; 1.8; и 2.9. Корень средний квадрат смещения (RMS) равен 2.25 %. Неопределенность относительно номинальных значений равна $u(Cref) = 1.5\%$.
4	Преобразовать компоненты в стандартные неопределенности $u(x)$	Доверительные интервалы подобных распределений должны быть преобразованы в стандартную неопределенность. $u(R_w) = 3.34/2 = 1.67\%$ $u(bias) = \sqrt{RMS_{bias}^2 + u(Cref)^2}$ $= \sqrt{2.25^2 + 1.5} = 2.71\%$
5	Расчитать суммарную стандартную неопределенность u_c	Стандартные неопределенности должны быть возведены в квадрат и просуммированы. Суммарная стандартная неопределенность равна квадратному корню из этой суммы квадратов. $u_c = \sqrt{u(R_w)^2 + (u(bias))^2} = \sqrt{1.67^2 + 2.71^2} = 3.18$
6	Расчитать расширенную Неопределенность $U = 2 \cdot u_c$	$U = 2 \cdot 3.18 = 6.36 \approx 6\%$



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Сколько значимых цифр?

Для измерений, проводимых посредством современных приборов, цифры обычно выводятся в электронном виде и должны использоваться как таковые даже если не понятно какие из них реально значимы. Только когда Вы пишете конечный результат, то число следует округлить до соответствующего числа значимых цифр.

Способ, которым записан результат, должен сказать кое-что о точности результата. Чем больше цифр приведено, тем более подразумевается точность. Значимые числа это те числа, которые содержат полезную информацию. Может быть не уместным приводить длину плавательного бассейна с точностью до долей миллиметра.

Для решения вопроса о том, что такое корректное число, необходимо знать неопределенность измерения. Это может быть следствием знаний, опыта или здравого смысла аналитика, или может быть получено из стандартного отклонения повторных экспериментов.

Если известно стандартное отклонение или 95% доверительный интервал, то записывайте эти числа с двумя значащими цифрами. Тогда результат измерения может быть записан с тем же количеством десятичных знаков. Например: $(1.123 \pm 0.032) \text{ M}$.

Является ли полученная неопределенность разумной?

На этот вопрос нет простого ответа! Все зависит от того для ответа на какой вопрос требуется Ваш результат и как много времени и ресурсов Вам выделено. Важно, что Вы должны выдать измерение с достаточной правильностью для того, чтобы дать возможность принять необходимое решение.

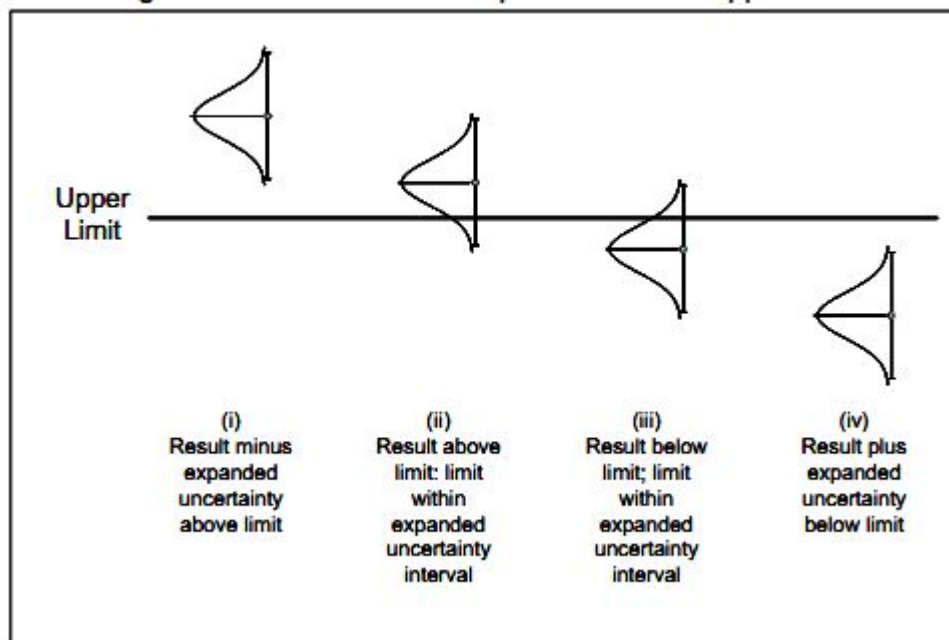


СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Чтобы дать ответ на этот вопрос следует воспользоваться руководством EURACHEM / CITAC Guide «Use of uncertainty information in compliance assessment», First Edition 2007.

Для принятия решения о соответствии полученного результата какой-либо спецификации, необходимо принять во внимание неопределенность измерения. На рисунке 1 показано типичные сценарии, которые возникают, когда результаты измерений, например концентрации аналита, используются для оценки соблюдения верхнего предела спецификации.

Figure 1 Assessment of Compliance with an Upper Limit



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Вертикальные линии показывают $\pm U$ расширенную неопределенность для каждого результата и связанные кривой указывают на выведенную плотность вероятности функции для значения, показывая большую вероятность значения, находится оно рядом с центром расширенной неопределенности интервала или вблизи концов. Случаи i) и iv) достаточно ясны; результаты измерений и их неопределенности обеспечивают хорошее доказательство того, что значение соответственно значительно выше или ниже предела. В случае (ii) однако, существует высокая вероятность того, что значение выше предела, но предел тем не менее находится в диапазоне интервала расширенной неопределенности.

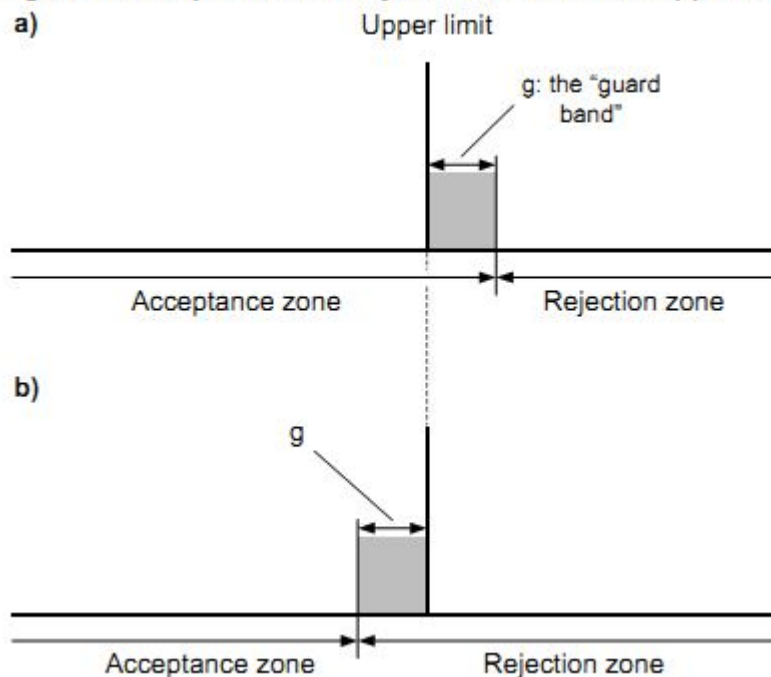
В зависимости от обстоятельств, особенно связанных с риском принять неправильное решение, вероятность принятия неправильного решения может быть минимизирована, чтобы оправдать решение о несоответствии результата заявленной спецификации. Аналогичным образом в случае (iii) вероятность того, что значение находится ниже предела может или не может быть достаточной, чтобы принять результат для оправдания соответствия спецификации.

Этот документ содержит дополнительные указания по настройке соответствующих критериев для принятия однозначных решений по вопросам соответствия, учитывая результаты связанные с информацией о неопределенности.

Ключом к оценке соответствия является принятие концепции «Правила разрешения». Эти правила дают рецепт для принятия или отклонения продукта, основанный на результате измерения, его неопределенности и пределов спецификации или ограничений, принимая во внимание приемлемый уровень вероятности принятия неправильного решения. На основе «Правила разрешения» определяются «зоны приемлемости» и «зоны исключения», таким образом, что если результат измерения находится в зоне приемлемости, то продукт объявляется соответствующим, а если в зоне исключения, то продукт признается несоответствующим.

СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Figure 2: Acceptance and Rejection zones for an Upper Limit



Начало зоны исключения находится в пределах спецификации L плюс значение g , (так называемой предохранительной полосы).

Значение g выбирается так, чтобы для измерения результата, большего или равного $L+g$ вероятность ложного отклонения была меньше или равна α . То есть, если результат находится в зоне исключения, то принятое правило дает низкую вероятность того чтобы разрешенный предел на самом деле не было превышено. (Рис. 2 а).

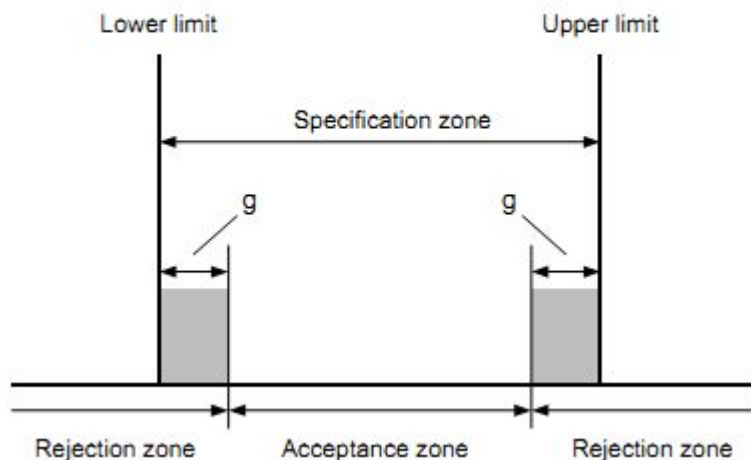
На (рис 2b) g был выбран для обеспечения низкого риска ложного признания.

В общем случае g будет кратно u стандартной неопределенности. Для случая, когда распределение вероятных значений приблизительно нормально, значение $1.64 \cdot u$ даст вероятность 5% , а значение $2.33 \cdot u$ означает вероятность 1%.



СОВРЕМЕННЫЕ МЕТОДЫ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ

Figure 3: Acceptance and rejection zones for simultaneous Upper and Lower Limits



В некоторых случаях спецификация устанавливает верхний и нижний пределы, например контроля состава.

Рис 3 показывает зоны принятия и исключения для такого случая, где предохранительные полосы выбраны так, что для образца существует высокая вероятность того, что параметр находится в пределах спецификаций.

Чем больше значение u тем больше доля образцов, которые будут оцениваться неправильно. Чем меньше значение u , тем выше будет стоимость анализа. Таким образом в идеале u должна быть выбрана оптимально для минимизации затрат на анализ, однако не в ущерб правильности.

Информация о том как это сделать однако очень редко доступна. В некоторых случаях, когда спецификации устанавливает верхний и нижний пределы, максимальный допустимый размер u определяется как доля разницы между этих пределов. Например одна такая спецификация утверждает, что расширенная неопределенность должна быть не более чем одной восьмой части этой разницы. Общий подход должен проводить скрининг измерений с помощью метода сравнительно дешевого и относительно большой неопределенности, а проверять с помощью метода с небольшой неопределенности для этих образцов, для которых в результате скрининга не было выработано четкого решения.

Как согласуется погрешность с неопределенностью?

Формулы пересчета параметров неопределенности в параметры погрешности :

$$\hat{S} = \hat{S}_{\Sigma} \cdot \sqrt[4]{\frac{n-1}{\hat{\nu}_{eff}}}$$

где

- \hat{S} - оценка СКО случайной погрешности
- \hat{S}_{Σ} - суммарная стандартная неопределенность
- N - количество результатов параллельных измерений
- $\hat{\nu}_{eff}$ - Оценка эффективного количества степеней свободы

$$\hat{S}_{\theta} = \sqrt{\hat{S}_{\Sigma}^2 - \hat{S}^2}$$

где

- \hat{S}_{θ} - оценка СКО неисключенной систематической погрешности (НСП)
- \hat{S} - оценка СКО случайной погрешности
- \hat{S}_{Σ} - суммарная стандартная неопределенность



Как согласуется погрешность с неопределенностью?

Формулы пересчета параметров неопределенности в параметры погрешности :

$$\theta_{(p)} = 1,1 \cdot \hat{S}_\theta$$

где
 \hat{S}_θ - оценка СКО неисключенной систематической погрешности (НСП)
 $\theta_{(p)}$ - оценка границы НСП

$$\Delta_{0,95} = \frac{t_{0,95}(n-1) \cdot \hat{S} + \hat{\theta}_{(p)} \cdot \hat{S}_\Sigma}{\hat{S} + \hat{S}_\theta} \cdot \hat{S}_\Sigma$$

Где
 $\Delta_{0,95}$ - оценка погрешности
 \hat{S} - оценка СКО случайной погрешности
 \hat{S}_Σ - суммарная стандартная неопределенность
 \hat{S}_θ - оценка СКО неисключенной систематической погрешности (НСП)
 $\theta_{(p)}$ - оценка границы НСП
 n - количество результатов параллельных измерений
 $t_{0,95}$ - коэффициент Стьюдента для (n-1) степеней свободы



Поэтапная валидация метода

Валидация - процесс установления характеристик и ограничений метода, идентификации влияний, которые могут изменять эти характеристики и в какой степени. С другой стороны, валидация - это процесс подтверждения того, что метод является пригодным для соответствующей цели, то есть, может быть использован для решения специфической аналитической проблемы.

В соответствии со стандартом ISO/IEC 17025 все методы испытания, используемые в лабораториях, должны быть валидированы и утверждены. Валидация методов выполняется для того, чтобы достичь необходимых характеристик типа точности, воспроизводимости и надежности результатов испытания.

ОБЩИЙ НАБОР ПОКАЗАТЕЛЕЙ ВАЛИДАЦИИ

- Специфичность/ селективность
- Правильность/точность (процент возврата Rec)
- Прецизионность (воспроизводимость S_R / сходимости S_r)
- Чувствительность I (LOD, LOQ)
- Линейность
- Диапазон измерения
- Робастность
- Неопределенность U

Поэтапная валидация метода

Основопологающим документом по проведению валидации аналитических методов является руководство ЕВРАХЕМ.

The Fitness for Purpose of Analytical Methods

EURACHEM Guide



The Fitness for Purpose of Analytical Methods

A Laboratory Guide to Method Validation
and Related Topics

Eurachem  *A FOCUS FOR
ANALYTICAL CHEMISTRY
IN EUROPE*

Поэтапная валидация метода

Важные характеристики валидации для различных типов ИСПЫТАНИЯ

Тип испытания / Требование	Идентификация	Тест на примеси		Содержание или эффективность действия
		Количественно	Предел	
Точность	-	+	-	+
Прецизионность:				
Сходимость	-	+	-	+
Воспроизводимость	-	- (1)	-	- (1)
Специфичность	+	+	+	+(2)
Предел обнаружения	-	+	+	-
Предел количественного определения	-	+	-	-
Линейность	-	+	-	+
Диапазон	-	+	-	+

- обычно не измеряется

+ обычно измеряется

(1) может быть необходим в некоторых случаях

(2) Может не требоваться в некоторых случаях

Поэтапная валидация метода



Специфичность (Specificity) (Селективность)

Руководство ICH Q2A использует термин «специфичность», несмотря на то что, например, в IUPAC в настоящее время обсуждается вопрос о замене этого термина на термин «селективность». В предварительных рекомендациях IUPAC «Селективность в аналитической химии» уточняется различие между обоими терминами следующим образом: «Специфичность является предельным случаем селективности».

Специфичность (specificity) — это способность методики безусловно определять анализируемое вещество в присутствии компонентов, которые могут быть в пробе. Как правило, ими являются примеси, продукты распада, плацебо и т. д. Отсутствие специфичности у отдельной аналитической методики может быть компенсировано использованием другой дополнительной аналитической методики.

Это определение имеет следующие приложения:

- проверка подлинности (обеспечить подлинность анализируемого вещества);
- испытания на чистоту (обеспечить тот факт, что выполненные аналитические методики позволяют правильно определить содержание примесей, т. е. родственных примесей, тяжелых металлов, остаточных растворителей и т. д.);
- количественное определение (обеспечить точность измерения при определении содержания/активности анализируемого вещества в пробе).

Идеальным случаем подтверждения специфичности (селективности) является отсутствие на хроматограммах посторонних пиков из плацебо или отсутствие интерференций пиков искомого аналита с пиками холостой пробы.

Поэтапная валидация метода

Например, в случае лекарственного препарата изготовленного из стандартизированного растительного экстракта специфичность аналитической методики подтверждается путем доказательства того, что:

- все вспомогательные вещества, используемые при изготовлении лекарственного препарата, при указанных условиях хроматографирования или не дают никакого пика, или дают пики, отчетливо отличающиеся от пиков определяемого вещества;
- действующее вещество, содержащееся в экстракте, четко определяется в присутствии в экстракте других веществ;
- подлинность определяемого действующего вещества доказана с помощью какого-либо стандартного образца с известной или доказанной структурой, путем сравнения как времен удерживания, так и спектров на различных длинах волн.

Оценка специфичности методики определения наличия определенных компонентов экстракта в лекарственном препарате проводится путем сравнения ВЭЖХ-хроматограммы с хроматограммой плацебо (смеси всех компонентов без экстракта) и/или аутентичного экстракта с доказанной подлинностью или смеси стандартных образцов. Примеры хроматограмм представлены далее.

Поэтапная валидация метода

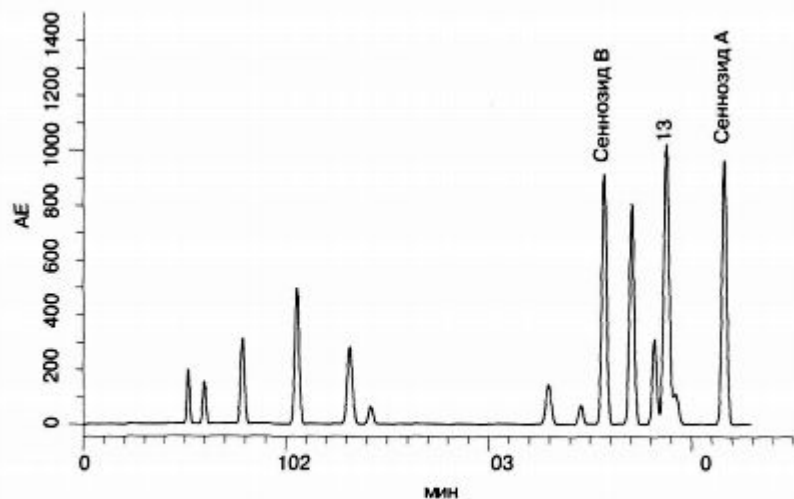


Рис. 1. ВЭЖХ-хроматограмма лекарственного препарата, содержащего в качестве действующего вещества экстракт плодов сенны (длина волны детектирования 270 нм)

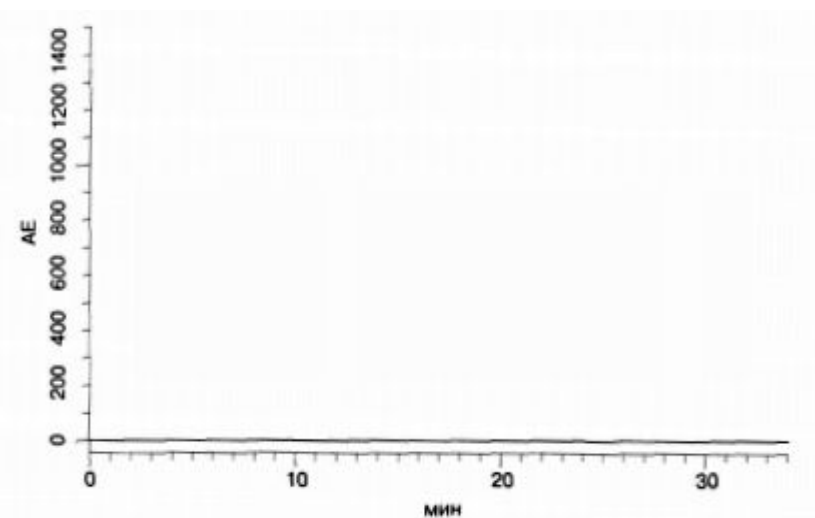


Рис. 2. ВЭЖХ-хроматограмма плацебо лекарственного препарата (длина волны детектирования 270 нм)

Поэтапная валидация метода

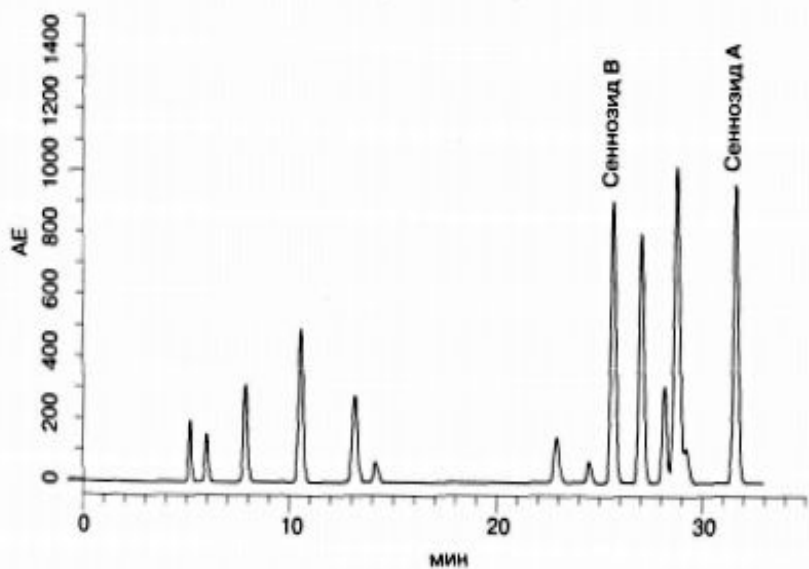


Рис. 3. ВЭЖХ-хроматограмма экстракта с доказанной подлинностью (длина волны детектирования 270 нм)

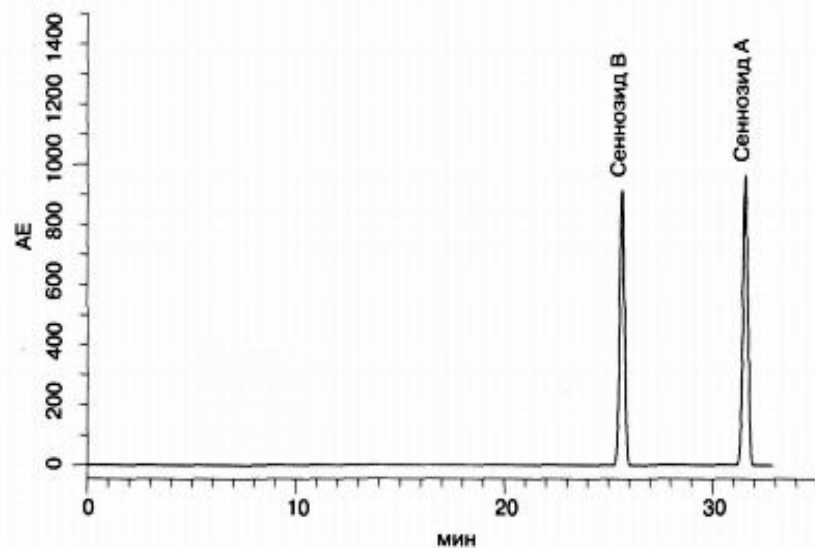


Рис. 4. ВЭЖХ-хроматограмма смеси стандартных образцов сеннозидов А и В (длина волны детектирования 270 нм)

Поэтапная валидация метода

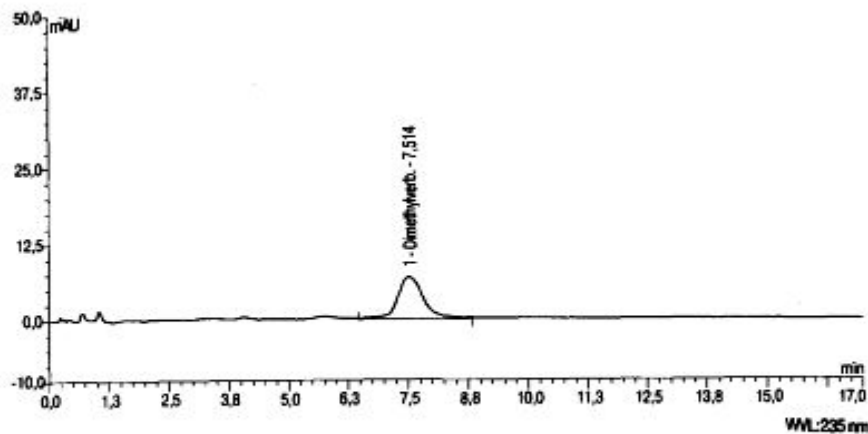
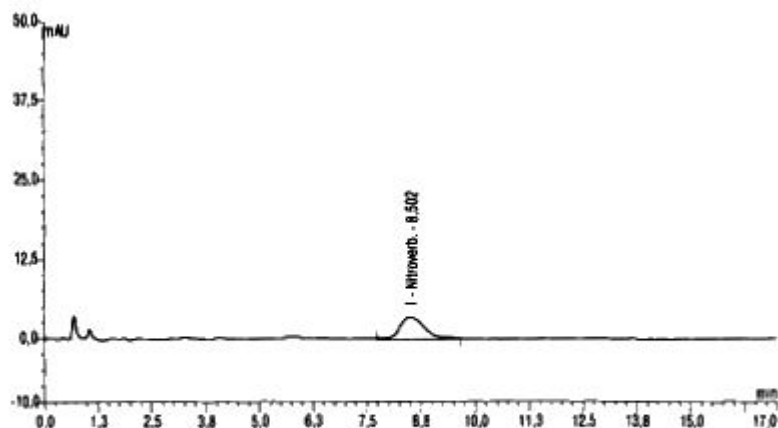
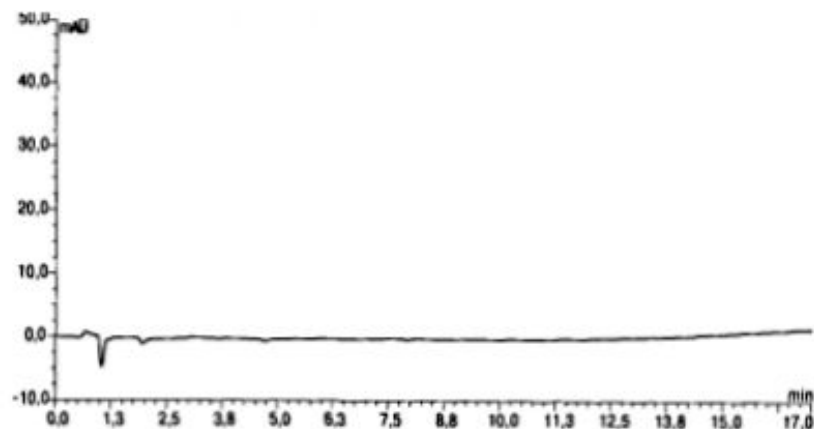
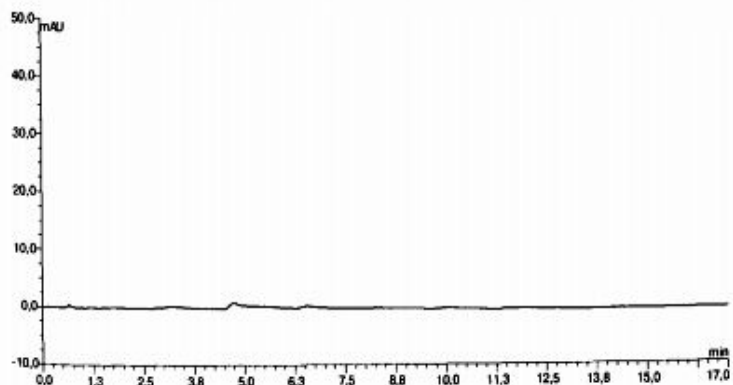
В случае лекарственного препарата изготовленного из субстанции, полученной химическим синтезом специфичность аналитической методики является доказаной, если ни используемый растворитель, ни подвижная фаза, ни компоненты матрикса (плацебо), ни примеси не искажают результат в пределах требуемой прецизионности.

Специфичность методики доказывается путем получения хроматограммы действующего вещества, растворителя, плацебо и (при необходимости) подвижной фазы, а также отдельных примесей. Должно быть показано, что на специфичность аналитической методики не влияет ни один из вышеназванных факторов.

Если, кроме того, эта методика также используется для исследований стабильности, необходимо проверить, не влияют ли на специфичность методики продукты деградации.

Например: Нитрендипин с тремя известными примесями.

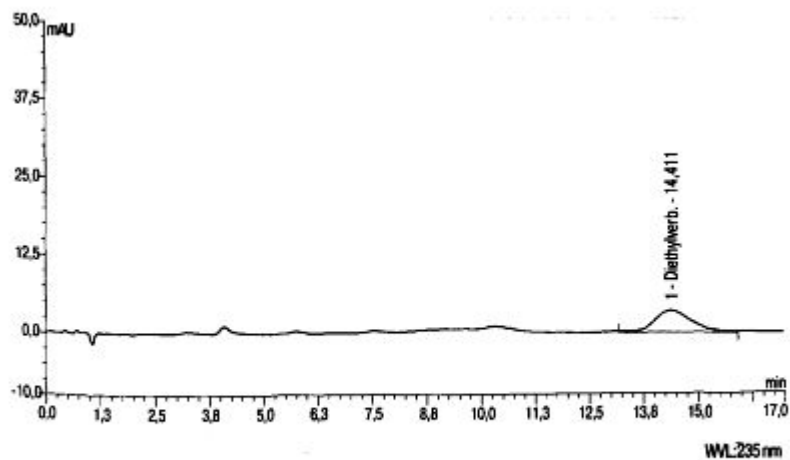
Поэтапная валидация метода



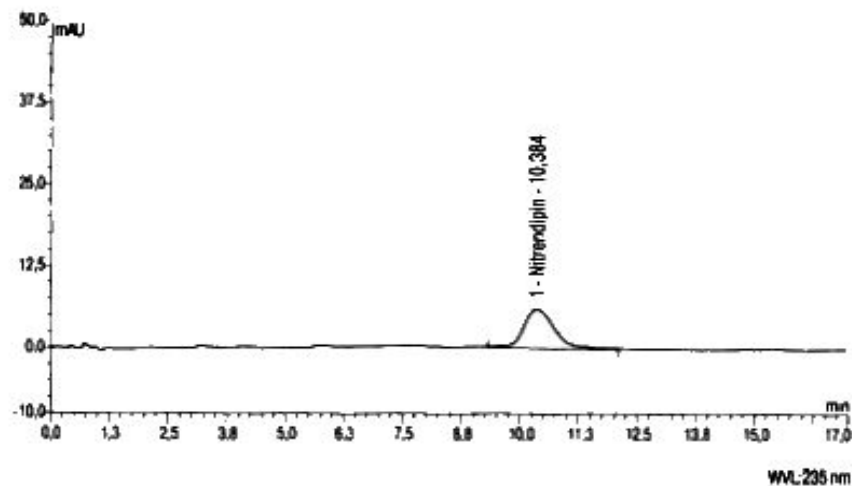
ВЭЖХ-хроматограмма нитропроизводного (примесь А: этил-метил-2,6-диметил-4-(3-нитрофенил)-1,4-пиридин-3,5-дикарбоксилат)

ВЭЖХ-хроматограмма диметилпроизводного (примесь В: диметил-2,6-диметил-4-(3-нитрофенил)-1,4-дигидропирдин-3,5-дикарбоксилат)

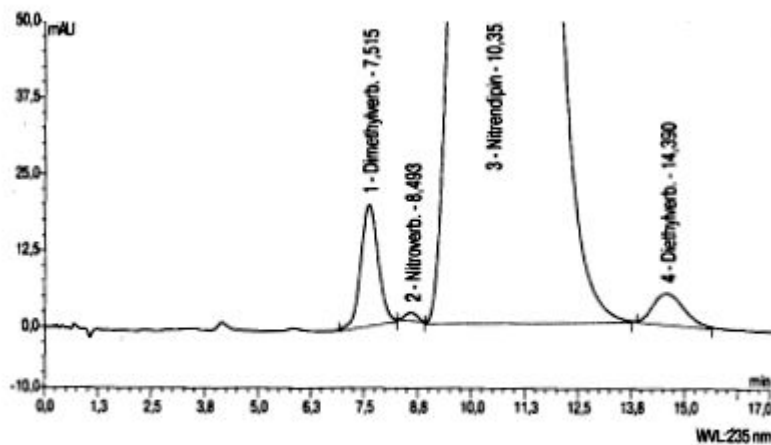
Поэтапная валидация метода



ВЭЖХ-хроматограмма диэтилпроизводного диметил-4-(3-нитрофенил)-1,4-дигидро



100%



ВЭЖХ-хроматограмма нитрендипина с примесями при их максимально допустимых концентрациях

Поэтапная валидация метода

Селективность в анализе может принимать две ступени :

Качественная — ступень, в которой другие вещества не влияют на определение искомым анализом согласно применяемой процедуре анализа образца.

Количественная — ступень, используемый в сочетании с другим словом (например постоянная, коэффициент, индекс, фактор, число) для количественной характеристики влияний (помех). Т.е. в случае интерферентных помех в определении искомого анализа необходимо введение корректировочных поправок в расчетные формулы, учитывающие влияние.

Согласно руководства ЕВРАХЕМ необходимую документацию, подтверждения селективности оформляют помимо приведенных выше хроматограмм, следующим образом:

Анализируют образец холостой пробы матрицы, как указано в описании анализа десятикратно .

Результат расчета площади в области выхода анализируемых пиков не должен отклоняться значительно от 0, т.е. фактически посторонние пики должны отсутствовать. Если все же имеется какой-либо интерферирующий компонент, то обязательно должна быть сделана коррекция нуля введением соответствующей поправки к вычислениям (вычитание площади интерферирующего компонента из площади анализируемого). Этот факт должен быть зафиксирован в методе.

Поэтапная валидация метода

Диапазон (Range).

Разработка метода измерения предполагает предварительный выбор рабочего диапазона. Рабочий диапазон зависит от:

а) Практических целей градуировки (калибровки).

Рабочий диапазон должен включать, насколько это возможно, весь диапазон для анализа различных образцов. Концентрация пробы, которая чаще всего встречается, должна находиться в центре рабочего диапазона.

б) Возможности технической реализации.

Получаемые значения измеряемых величин, нужно линейно соотносить с концентрациями. Это налагает некоторые требования, например, чтобы значение измеряемых величин, в нижнем пределе рабочего диапазона значительно отличались бы от значений холостого опыта. Поэтому нижний предел рабочего диапазона должен быть равен или больше предела определения данной методики. Разводить или концентрировать пробу нужно без риска систематической ошибки.

в) Дисперсия значений измеряемых величин, не должна зависеть от концентрации.

Эту независимость проверяют статистическим тестом. Суть теста заключается в проверке однородности дисперсий на границах рабочего диапазона. Данные для этого теста получают следующим образом :

Поэтапная валидация метода

С целью проверки однородности дисперсий десять раз параллельно измеряют для концентраций на нижнем и верхнем пределах (x_{min} и x_{max}) рабочего диапазона. В результате этих измерений получают десять значений измеренных величин, y_{ij} .

Оба ряда данных с концентрациями x_{min} и x_{max} используют для расчета дисперсий и по уравнению :

$$S_i^2 = \frac{\sum_{j=1}^{10} (y_{i,j} - \bar{y}_i)^2}{n_i - 1}$$

Дисперсии с помощью F-критерия проверяют на значимость расхождений на пределах рабочего диапазона . Для этого вычисляют с помощью уравнения значение PG:

$$PG = \frac{S_{max}^2}{S_{min}^2}, \text{ если } S_{max}^2 > S_{min}^2$$

$$PG = \frac{S_{min}^2}{S_{max}^2}, \text{ если } S_{min}^2 > S_{max}^2$$

PG сравнивают с табличными значениями квантиля F-распределения.

- a) если $PG < F_{f1, f2, 0.99}$ разница между дисперсиями незначима
- b) если $PG > F_{f1, f2, 0.99}$ разница между дисперсиями значима

Если разница между дисперсиями значима, то стоит уменьшить рабочий диапазон до размеров, при которых расхождение дисперсий будет случайно.

Поэтапная валидация метода



Точность (Accuracy) и правильность (Trueness)

Точность — близость соглашения между результатом испытания и принятым референтным (приписанным) значением. Термин точность при применении к множеству (выборке) результатов испытаний включает комбинацию компонентов случайной ошибки и общей систематической ошибки или компонента смещения. Иногда употребляется термин «истинность» (trueness).

Точностью также называют количественную разницу между средним из набора результатов или индивидуального результата и значения, которое принято как истинное или правильное значение для измеренного количества. Степень соответствия между аналитическим результатом и истинным значением должна быть высокой.

Правильность аналитической методики количественного определения доказывается на всем диапазоне применения. Для этого может быть использован один из представленных далее способов.

В качестве приводимых данных для подтверждения точности обычно приводят графические данные о регрессионной зависимости коэффициента возврата (Recovery) от анализируемой концентрации и данные о распределении так называемых остатков (разниц между фактическими значениями и значениями, полученными из регрессионной зависимости). Указывают также коэффициенты возврата, RSD для них.

Поэтапная валидация метода

1. Метод с плацебо с нулевым содержанием определяемого компонента.

В методе с плацебо определяется фактор отклика действующего вещества в модельной смеси компонентов плацебо. Концентрация компонентов плацебо составляет 100% навески от номинального, указанной в методике контроля. Действующее вещество добавляется в модельную смесь в соответствии с требуемым уровнем концентрации (например, 80,100 и 120% содержания субстанции в лекарственном препарате).

2. Метод добавок

В методе добавок фактор отклика определяется путем добавления действующего вещества к образцу с известной концентрацией субстанции. Например, навеска содержит 70% целевой концентрации в пробе; затем добавками содержание субстанции доводят до 80,100,120%.

Для определения правильности при обоих методах (метод с плацебо и метод добавок) проводится как минимум 9 испытаний с не менее 3 концентрациями в определенном диапазоне применения методики (например, 3 концентрации с 3-кратным определением для каждой концентрации, с выполнением всех стадий аналитической методики).

Оценка проводится путем расчета процента нахождения известного добавленного количества действующего вещества, стандартного отклонения, коэффициента вариации (CV) и доверительного интервала среднего значения (P - 95%).

Поэтапная валидация метода

3. Метод сравнения

Сравнение результатов аналитической методики с результатами второй, независимой и провалидированной методики. Расчет проводится с учетом статистических тестов (например, F- и t-критерии).

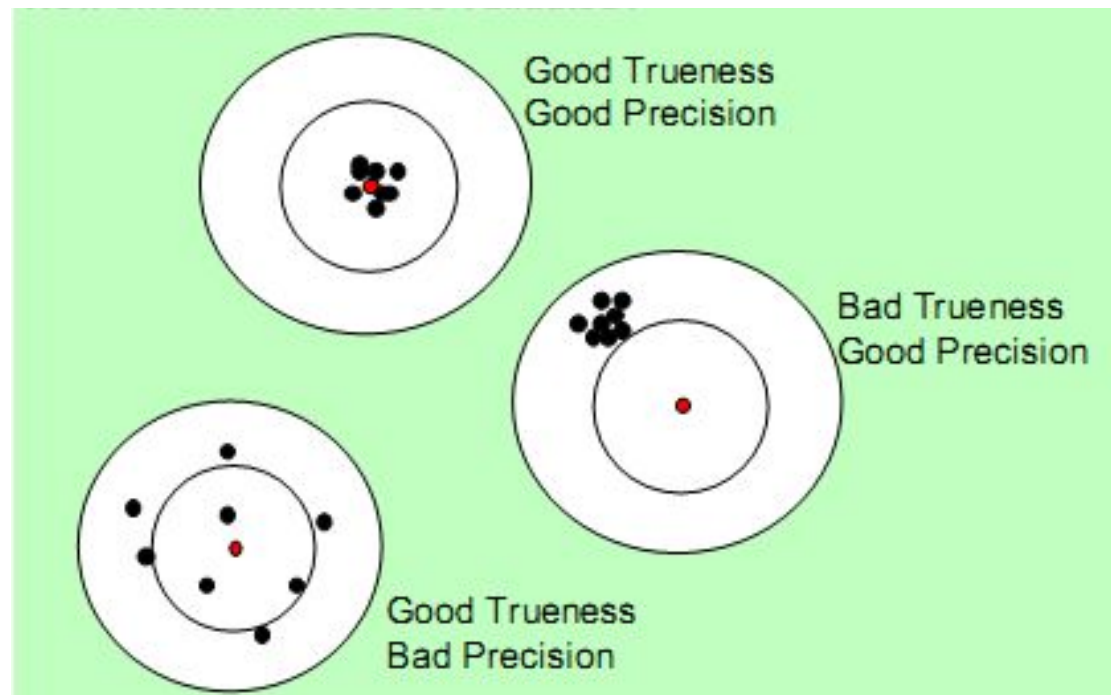
Для оценивания точности (правильности) используют различные критерии:

- 1) отношение разности величин (заданной и истинной) к заданной величине. Если процентное отклонение от заданной величины меньше $3S$, правильность результатов находится в допустимых пределах;
- 2) по допустимому отклику в зависимости от концентрации активного компонента (см. таблицу), для биоаналитических методов – не более 15 %, за исключением предела количественного определения – не более 20 %.

Оценка правильности в зависимости от концентрации активного компонента в анализируемом образце

Активный компонент, %	Коэффициент возврата, %
100	98 – 102
≥ 10	98 – 102
≥ 1	97 – 103
$\geq 0,1$	95 – 105
0,01	90 – 107
0,001	80 – 110
0,0001	80 – 110
0,00001	60 – 115
0,000001	40 – 120

Поэтапная валидация метода



Поэтапная валидация метода



Прецизионность (precision).

Прецизионность аналитической методики выражает степень близости (степень дисперсии) между сериями измерений, полученных при параллельных измерениях одного однородного образца, в установленных условиях.

Прецизионность может исследоваться на следующих уровнях: повторяемость, промежуточная прецизионность и воспроизводимость. Прецизионность должна оцениваться на однородных аутентичных образцах. Однако если невозможно получить однородный образец, она может определяться с использованием специально приготовленных образцов или испытуемых растворов. Прецизионность аналитической методики обычно выражается как дисперсия, стандартное отклонение или коэффициент вариабельности серий результатов измерений.

Прецизионность зависит только от распределения случайной ошибки и не касается истинного значения или приписанного значения. Количественные меры точности критически зависят от предусмотренных условий. Сходимость (повторяемость) и воспроизводимость — специфические наборы чрезвычайных условий. *[Прим. к рус. переводу. Английский термин «precision», характеризующий в аналитических измерениях степень рассеяния результатов при повторении процедуры анализа одного и того же образца, не имеет общепринятого русского эквивалента. Если принять для него термин “воспроизводимость” (Рекомендации научного Совета по аналитической химии АН СССР, 1975), это внесет путаницу в перевод подтерминов “repeatability” («сходимость») и “reproducibility” (“воспроизводимость”), отвечающих соответственно минимальному и максимальному варьированию влияющих факторов.*

Поэтапная валидация метода

Во избежание недоразумений, слово “точность”, которое иногда используется для перевода «precision», должно относиться исключительно к термину “accuracy”, характеризующему степень близости результата к истинному значению измеряемой величины. Поэтому представляется оправданным использование для «precision» несколько непривычного термина «прецизионность». Наиболее близким по значению может быть слово «кучность».]



Сходимость (повторяемость)

Сходимость должна показать, что методика анализа при проведении в одинаковых условиях обеспечивает получение сравнимых результатов.

Аналитические условия должны быть следующие.

- тот же самый лаборант
- тот же самый метод анализа
- тот же самый образец
- та же самая аппаратура
- в тот же самый день
- одни и те же реактивы

Сходимость может быть показана при помощи:

- анализа не менее 6 подготовленных проб при 100% концентрации (номинальном содержании) фармацевтической субстанции в исследуемом растворе;
- анализа не менее 9 подготовленных проб в рамках диапазона применения этой методики (например, 3 концентрации, 3 повтора).

Поэтапная валидация метода

Оценка и расчет результатов проводится путем вычисления среднего значения, стандартного отклонения, коэффициента вариации и доверительного интервала. Критерий приемлемости для коэффициента вариации необходимо устанавливать в зависимости от предполагаемых возможных значений (например, границы нормы для содержания исходного вещества в спецификации). В зависимости от определенного коэффициента вариации и границ нормы в спецификации устанавливается число испытаний, необходимых для рутинного контроля каждой пробы.

Таблица 1. Отношение между количеством анализов и коэффициентом вариации для определения достоверного отклонения от заданного значения, $p = 95\%$ [по Гриму (1973)]

Число испытаний, <i>n</i>	Коэффициент вариации								
	0,1	0,2	0,3	0,4	0,5	1	1,5	2	2,5
1	0,9	1,8	2,70	3,60	4,5	9,0	13,5	18,0	22,5
2	0,25	0,50	0,74	0,99	1,2	2,5	3,7	5	6,2
3	0,16	0,32	0,48	0,64	0,8	1,6	2,4	3,2	4
4	0,12	0,25	0,37	0,5	0,6	1,2	1,9	2,5	3,1
5	0,10	0,21	0,31	0,42	0,5	1	1,6	2,1	2,6
6	0,09	0,19	0,28	0,37	0,5	0,9	1,4	1	2,3
7	0,08	0,17	0,25	0,34	0,4	0,8	1,3	1,7	2,1
8	0,08	0,15	0,23	0,31	0,4	0,8	1,2	1,5	1,9
9	0,07	0,14	0,21	0,29	0,4	0,7	1,1	1,4	1,8

Поэтапная валидация метода

Пример: Если нормы количественного содержания действующего вещества в фармацевтической субстанции составляют $100 \pm 2\%$ и коэффициент вариации аналитической методики составляет 1,5%, то при количестве испытаний $n = 3$, не может быть принято статистически достоверное заключение о соответствии спецификации ($\pm 2,4\%$ статистически достоверное). Скорее всего, для этого должно быть проведено не менее $n = 4$ испытаний ($\pm 1,9\%$ статистической достоверности).

Не имеется никаких официальных требований для максимального размера стандартного отклонения и относительного стандартного отклонения (s_r) (или коэффициента вариации (CV), что фактически есть суть одно и то же). Зависит от типа анализа. Стандартное отклонение может изменяться между 2 и 20 %. Для анализа лекарственных соединений при контроле качества стандартное отклонение – менее 5 %. Для биологических и биоаналитических методов – не более 15 % для каждого концентрационного уровня. Исключение составляет определение остаточных количеств веществ, которые балансируют на пределе количественного определения, поэтому для этого случая – не более 20 %.

Предел сходимости « r »: значение меньше чем или равное тому, для которого абсолютная разность между двумя результатами испытания, полученными в условиях сходимости, можно ожидать нахождения в пределах вероятности 95 %.

Сходимость (предел) вычисляют по формуле:

$$r = 2.8 * CV \quad \text{или} \quad r = 2.8 * S_r$$

Поэтапная валидация метода

Принятые критерии для точности во многом зависят от типа анализа. Американской ассоциацией аналитической химии (АОАС) утверждена программа для аналитических методов, в которой приводятся данные оценки сходимости как функции аналитических концентраций.

Данные оценки сходимости как функции концентрации анализируемого компонента

Содержание анализируемого компонента, %	Относительное стандартное отклонение (RSD)
100	1,3
10	2,8
1	2,7
0,1	3,7
0,01	5,3
0,001	7,3
0,0001	11,0
0,00001	15,0
0,000001	21,0
0,0000001	30,0

Поэтапная валидация метода

Для оценивания допустимых пределов сходимости и воспроизводимости иногда прибегают к так называемым критериям Горвица. Это расчетный критерий, устанавливающий эмпирическую зависимость между RSD (относительным СКО) и концентрацией вещества (C) в анализируемом объекте.

$$RSD_r = 0,67 \cdot 2^{(1-0,5 \cdot \log(c))}$$

Критерий Горвица для расчета
предела сходимости

$$RSD_R = 2^{(1-0,5 \cdot \log(c))}$$

Критерий Горвица для расчета
предела воспроизводимости

Поэтапная валидация метода



Воспроизводимость

Воспроизводимость - это прецизионность в условиях, когда результаты испытания получены одним и тем же самым методом на идентичных испытательных образцах в различных лабораториях с различными операторами, использующими различное оборудование.

Пределом воспроизводимости «R» является значение меньшее либо равное тому, для которого абсолютная разность между двумя результатами испытания, полученными в условиях воспроизводимости, находится в пределах вероятности 95 %.

Воспроизводимость (предел) вычисляется по формуле: $R = 2.8 * s_R$

Метод должен быть стабильным, то есть, характеристики должны быть воспроизводимыми при повторении от дня к дню, от лаборанта к лаборанту, от лаборатории к лаборатории, от аппарата к аппарату и т.д.

В целом воспроизводимость оценивается как степень разброса результатов, получаемых анализируемым методом. Статистическими показателями разброса результатов являются среднеквадратичное отклонение и относительный показатель разброса результатов – коэффициент вариации CV. Как правило, частным применением исследования воспроизводимости является промежуточная прецизионность. Обычно этот критерий валидации рассчитывают в пределах одной лаборатории для разных операторов или разных приборов. Воспроизводимость – понятие, предполагающее проведение более широкомасштабных межлабораторных испытаний, что не всегда доступно.

Поэтапная валидация метода



Линейность

Линейность аналитической методики доказана, если в выбранном диапазоне применения можно показать прямо пропорциональное соотношение между концентрацией исследуемого вещества в растворе и сигналом детектора (фактором отклика).

Линейность методики подтверждается в диапазоне концентраций от 80 до 120% (от номинала) действующего вещества в исследуемом растворе (на ряде разведений исходного раствора). При этом измеряется и анализируется как минимум 5 концентраций, причем значение для каждого анализа получается путем расчета среднего значения из нескольких инъекций (вколов).

Оценка проводится по наличию линейной зависимости между концентрацией вещества и сигналом с помощью подходящего регрессионного уравнения (например, полученного методом наименьших квадратов). Характеристики линейности методики выражаются в виде коэффициента корреляции, наклона прямых, а также величины отрезка на оси ординат. В качестве дополнительной информации для оценки линейности может быть приведено графическое представление и отклонение полученных значений от регрессионной прямой.

Если для рутинного контроля показателя «количественное определение» предусматривается калибровка по одной точке, то доверительный диапазон отрезка на оси ординат должен включать нулевую точку. В этом случае необходимо оценивать линейность методики в диапазоне концентраций от 10 до 120% (от номинала). Наряду с линейной функцией для определенных типов детекторов может быть показана другая математическая зависимость, причем число калибровочных точек для рутинного контроля должно быть подобрано соответствующим образом.

Поэтапная валидация метода

Основные и статистически обоснованные принципы построения и оценивания калибровочных регрессионных зависимостей изложено в :

1. ДСТУ ISO 8466-1-2001 «ОПРЕДЕЛЕНИЕ ГРАДУИРОВОЧНОЙ ХАРАКТЕРИСТИКИ МЕТОДИК КОЛИЧЕСТВЕННОГО ХИМИЧЕСКОГО АНАЛИЗА. Часть 1. Статистическое оценивание линейной градуировочной характеристики»
2. ДСТУ ISO 8466-2-2001 «ОПРЕДЕЛЕНИЕ ГРАДУИРОВОЧНОЙ ХАРАКТЕРИСТИКИ МЕТОДИК КОЛИЧЕСТВЕННОГО ХИМИЧЕСКОГО АНАЛИЗА. Часть 1. Принцип оценивания нелинейной градуировочной характеристики второго порядка».
3. ДСТУ ISO 11843-2:2004 Способность к обнаружению. Часть 2. Методология в случае линейной калибровки).

Критерии, которые обязательно должны приводиться совместно с калибровочным графиком:

b	наклон калибровочного графика
a	пересечение графика с осью Y
r^2	коэффициент корреляции
S_y	остаточное среднеквадратичное отклонение
PG	критерий линейности из теста Манделя

Поэтапная валидация метода

Тест Мандела. Во время статистической проверки на линейность, градуировочные данные используют для расчета линейной и нелинейной градуировочной характеристики с остаточным средним квадратичным отклонением S_{y1} или S_{y2} .

$$S_{y1} = \sqrt{\frac{\sum_{i=1}^N [y_i - (a + b \cdot x_i)]^2}{N - 2}}$$

$$S_{y2} = \sqrt{\frac{\sum_{i=1}^N y_i^2 - a \cdot \sum_{i=1}^N y_i - b \cdot \sum_{i=1}^N x_i \cdot y_i - c \cdot \sum_{i=1}^N x_i^2 \cdot y_i}{N - 3}}$$

Расхождение дисперсий DS^2 рассчитывают по уравнению : $DS^2 = (N-2) \cdot S_{y1}^2 - (N-3) \cdot S_{y2}^2$
 DS^2 и дисперсию нелинейной градуировочной характеристики S_{y2} проверяют с помощью F- критерия (критерий Фишера для $F_{(0.01,1,N-3)}$), для выявления значимости расхождений. Для этого по уравнению рассчитывают критерий Мандела :

$$PG = \frac{DS^2}{S_{y2}^2}$$

Если $PG \leq F$: Нелинейная градуировочная характеристика не обеспечивает ощутимо лучшей аппроксимации, - значит, градуировочная характеристика линейна.

Если $PG > F$: Рабочий диапазон стоит уменьшить так, чтобы получить для него линейную градуировочную характеристику. Если это невозможно, значение измеряемых величин, для проанализированных проб нужно оценивать с применением нелинейной градуировочной характеристики.

Поэтапная валидация метода



Предел обнаружения (detection limit)

Предел обнаружения отдельной аналитической методики представляет собой наименьшее количество анализируемого вещества в пробе, которое может быть обнаружено, но не обязательно выражено точным значением. Для инструментальных методов определяют отношение сигнал/шум или измеряют величину фоновый сигнала. Для неинструментальных методов предел обнаружения определяется анализом образцов с известными концентрациями аналита и установлением минимального уровня аналита, при котором он может быть достоверно обнаружен. Согласно фармакопейной документации, установленный предел обнаружения и определения должен быть подтвержден анализом модельных смесей, содержащих аналит на пределе и близко к нему.



Предел количественного обнаружения (quantitation limit)

Предел количественного обнаружения отдельной аналитической методики представляет собой наименьшее количество анализируемого вещества в пробе, которое может быть оценено количественно с приемлемой правильностью и прецизионностью. Предел количественного обнаружения является параметром методик определения количественного содержания компонентов, присутствующих в образце плацебо (т. е. все вещества, кроме действующего вещества) в низких концентрациях, и в частности, используется для характеристики методик определения примесей и/или продуктов деградации.

Поэтапная валидация метода

Предел обнаружения (детектирования) для аналитического прибора LOD

Предел обнаружения (детектирования) прибора — самое низкое количество (выраженное как масса), которое может быть измерено с разумной статистической достоверностью. Обычно LOD определяют на стандартных растворах.

В классике жанра хроматографического метода, предел детектирования выражается как количество образца, соответствующее удвоенному или утроенному шуму анализа. Необходимость выполнения надлежащего количественного определения не является необходимым требованием при выполнении макроанализов.

Классическая формула для хроматографических методов для LOD :

$$LOD = \frac{k \cdot C}{(S / N)}$$

k	— коэффициент (обычно он равен от 2 до 3),
C	— концентрация,
S	— значение сигнала (высота пика),
N	— величина шума.
S/N	— пиковое соотношение сигнал/шум.

Поэтапная валидация метода

Для нехроматографических методов предел детектирования, выраженный как концентрация или количество, получают из наименьшей меры x_L , которая может быть обнаружена с разумной достоверностью для данной аналитической процедуры. Значение x_L рассчитывают по уравнению:

$$x_L = x_{bl} + k * s_{bl}$$

где

x_{bl}

— среднее холостого определения,

k

— численный фактор, выбранный согласно желательному уровню достоверности,

s_{bl}

— стандартное отклонение холостых измерений.

Предел обнаружения LOD и предел определения LOQ могут быть вычислены и расчетным методом

$$LOD = \frac{3,3 \cdot S_y}{b}$$

$$LOQ = \frac{10 \cdot S_y}{b}$$

где LOD – предел обнаружения, LOQ – предел количественного определения, S_y – остаточное стандартное отклонение регрессионной зависимости, b – наклон калибровочной кривой.

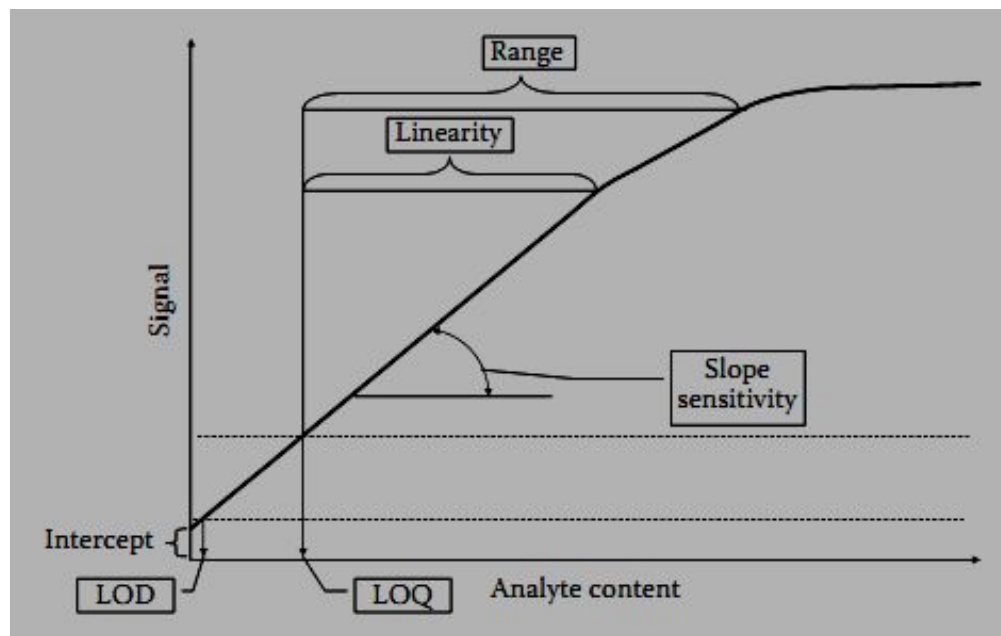
Уравнение для нахождения предела обнаружения имеет то преимущество, что оно получается полностью из калибровочного уравнения. Более статистически «основательно» уравнение получаемое из данных калибровки, которое опубликовано ISO (ISO 11843-2:2000) (наш эквивалент ДСТУ ISO 11843-2:2004 « СТАТИСТИЧЕСКИЙ КОНТРОЛЬ. СПОСОБНОСТЬ К ОБНАРУЖЕНИЮ.

Часть 2. Методология в случае линейной калибровки»):

Поэтапная валидация метода

$$LOD = \frac{2 \cdot t_{(0.05, n-2)} \cdot S_y}{b} \cdot \sqrt{\frac{1}{K} + \frac{1}{I \cdot J} + \frac{\bar{x}^{-2}}{J \cdot \sum_i (x_i - \bar{x})^2}}$$

Здесь калибровка выполняется с I независимыми калибровочными растворами (включая холостую пробу, если возможно то раствор со значением возле предполагаемого предела детектирования), каждый измеряется J раз. K представляет число повторов, которые должны проводиться для каждого испытуемого раствора для получения усредненного отклика. Чем больше повторов Вы можете проводить, тем ниже уровень аналита Вы сможете «поймать».



Поэтапная валидация метода



Возврат (Recovery)

Возврат (также употребляются термины «процентная мера правильности», «открываемость», «извлечение», «экстрагируемость».): частное от деления количества аналита, добавленного к образцу для испытаний (укрепленному или пробе со стандартной добавкой) до анализа, к разности результатов для укрепленного и неукрепленного образца, и выраженное в процентах; процент возврата (% Rec) рассчитывают следующим образом:

$$\text{Rec} = \frac{\text{количество аналита, добавленного к образцу для испытаний}}{\text{разность результатов для укрепленного и неукрепленного образца}} \times 100 \%$$

Процент возврата для анализа должен быть высоким и постоянным от анализа к анализу.

Что касается с медицинских препаратов, определение возврата должно выполняться с использованием образцов плацебо (холостых проб). В других случаях имеется свободный выбор между добавкой к плацебо (холостым пробам) или стандартными добавками.

Поэтапная валидация метода

Определение возврата добавкой активного ингредиента к плацебо (холостой пробе)

Прибавьте точно установленное количество активного ингредиента (аналита) к образцу плацебо (холостой пробе). Приготовьте образец, как изложено в обычной методике анализа. Сделайте 10 одиночных определений.

Проделайте определение возврата, добавляя около 80 %, 100 % и около 120 % ожидаемого активного содержания ингредиента (аналита) к плацебо(холостой пробе).

Особенно в случае продуктов разложения, которые образуются из активных ингредиентов (аналитов), определения возврата должны выполняться с добавкой 10 %, 50 % и 100 % ожидаемого максимального допустимого содержания, если это позволяет более низкий предел обнаружения (предел детектирования).

Анализ приемлем, если процент возврата находится между 97 и 103 % для макроопределений. Согласно Постановлению № 20 от 20.04.1999 г. Главного санитарного врача при определении содержания химических соединений в сельскохозяйственном сырье, продуктах питания и объектах окружающей среды процент возврата должен находиться в пределах от 70 до 120 % при ОСКО возврата не более 20 %. Уровень содержаний при этом предполагается от 0,01 до 0,5 мг/кг или мг/л. Если, доверительный интервал для процента возврата не включает 100%, то фактический результат должен быть исправлен с процентом возврата, т.е. коэффициент возврата должен быть задействован в формуле расчета конечного результата.

Поэтапная валидация метода

Определение возврата со стандартной добавкой

Сделайте гомогенный образец плацебо (холостой пробы). Определите содержание активного компонента (аналита) в объединенном одиночном анализе. Добавьте между 50 и 100 % количества активного ингредиента (аналита), обычно находящегося в матрице. Если уместно, можно добавить и меньшие количества (10 или 20 %) в зависимости от условий линейности и других условий анализа. Выполните 10 определений возврата.

$$\text{Rec} = \frac{I_a - I_p}{I_t} \times 100\%$$

I_a = фактическое количество, найденное при анализе

I_p = расчетное содержание в образце (образец плацебо, холостой пробы))

I_t = добавленное количество к образцу

Анализ приемлем, если процент возврата попадает между 90 и 110 % (При определении следов тяжелых металлов, пестицидов и других остатков общепринятые международные соглашения допускают более широкие границы для возврата 70–130 % и даже 60–140 %. Это зависит от содержаний и обычно регламентируется общими операционными процедурами, либо отражается в конкретных методиках анализа. Фактический результат должен быть исправлен с учетом процента возврата (что достигается введением множителя $1/\text{Rec}$ в расчетную формулу.

Поэтапная валидация метода



Робастность

Испытание на стабильность: внутрилабораторное исследование для изучения поведения аналитического процесса, когда производятся незначительные изменения в окружающей среде и/или эксплуатационных режимах.

Робастность (ошибкоустойчивость) аналитической процедуры является мерой ее способности оставаться незатронутой незначительными, но преднамеренными изменениями в параметрах метода и обеспечивает показателем ее надежности в течение нормального использования. (Заметим, что стабильность и робастность (ошибкоустойчивость) — синонимы.)

Определяется проведением теста правильности при небольших реальных отклонениях параметров методики или свойств анализируемого объекта. Процедура изучения стабильности должна показать стабильность аналитов в течение получения, обработки образца, длительного (замороженного при предписанной температуре хранения), кратковременного хранения (при температуре $(20 \pm 2)^\circ\text{C}$ от 4 до 24 ч), после циклов замораживания/оттаивания (три цикла) и самого аналитического процесса.

К процедуре исследования робастности метода подходят индивидуально в каждом отдельном случае. Например, метод ВЭЖХ предполагает исследование влияния отклонений содержания ингредиентов подвижной фазы от $\pm 5\%$ до $\pm 15\%$, изменения pH подвижной фазы на $\pm 0,25$ ед., изменения температуры колонки на $\pm 5^\circ\text{C}$, реакцию на смену колонок. Обычно при этом фиксируются данные по изменению RT или RTT, RSD площадей пиков, Rs для критичных пар пиков, асимметричность пиков, эффективность. В таких исследованиях рождается истина о приемлемости тех или иных условий, которые впоследствии и формируют критерии приемлемости хроматографической системы.

Поэтапная валидация метода

Этап валидации	Как выполнить ?	Что должно быть задокументировано ?
Селективность (специфичность)	<ol style="list-style-type: none"> 1. Проанализировать экстракт (раствор) плацебо или холостой пробы . (В случае отсутствия интерференций $n=5$, в случае наличия $n=10$) 2. Стандартные растворы примесей и определяемого аналита. 	<ol style="list-style-type: none"> 1. Хроматограммы плацебо. 2. Хроматограммы растворов , доказывающие отсутствие интерференций. 3. В случае наличия интерференций необходимо рассчитать СКО мешающего фактора , чтобы учесть его при расчетах.
Диапазон	<p>Проанализировать по 5 -10 повторностей растворов , предназначенных для построения калибровки на границах выбранного рабочего диапазона</p>	<ol style="list-style-type: none"> 1. Привести расчетный критерий (PG) 2. Привести дисперсии для границ диапазона. 3. Привести значение критерия Фишера для вероятности 0,99 и df_{min} df_{max}
Прецизионность => Сходимость	<p>В одних и тех же условиях , руками одного аналитика, проанализировать 10 повторностей приготовленного раствора аналита.</p>	<ol style="list-style-type: none"> 1. Коэффициент вариации CV 2. Предел сходимости r 3. Привести хроматограммы

Поэтапная валидация метода

Этап валидации	Как выполнить ?	Что должно быть задокументировано ?
<p><i>Прецизионность</i> => Промежуточная прецизионность</p>	<p>В одних и тех же условиях , руками разных аналитиков или на двух разных приборах, проанализировать по 10 повторностей приготовленного раствора аналита.</p>	<ol style="list-style-type: none">1. Коэффициент вариации CV в масштабах обеих условий.2. Предел воспроизводимости R3. Привести хроматограммы.
<p><i>Точность</i> => Правильность</p>	<p>Выполнить определение коэффициента возврата :</p> <p><u>1. В случае макроисследований :</u> В плацебо добавить аналит в известной концентрации, как правило в диапазоне от 50 до 150 % от номинала с шагом в 10-20 %. Приготовить примерно 15 образцов (5 концентраций по 3 повтора).</p> <p><u>2. В случае микроисследований :</u> В плацебо добавить аналит в известной концентрации, в диапазоне от 25 до 150 % от номинального содержания примеси.</p>	<ol style="list-style-type: none">1. Хроматограммы растворов.2. Рассчитать СКО.3. Рассчитать коэффициент рикавери для всех точек, а также среднее значение.4. Привести график , подтверждающий линейность рикавери на всем протяжении рабочего диапазона.5. Привести график распределения остатков.

Поэтапная валидация метода

Этап валидации	Как выполнить ?	Что должно быть задокументировано ?
Калибровка Линейность	Проанализировать стандартные растворы определяемого аналита, различных концентраций (желателен диапазон от 10 % до 120 % содержания от номинального значения). Количество калибровочных точек не менее 6 в тройной повторности .	<ol style="list-style-type: none">1. Привести хроматограммы2. Построить калибровку, указать коэффициенты регрессии (b, a).3. Рассчитать ошибку аппроксимации или остаточное СКО регрессии S_y4. Привести критерий линейности (PG) (Манделъ-тест) и соответственно табличный критерий Фишера $F_{(0.01, 1, N-3)}$5. Привести графики зависимости и остатков.
Предел детектирования Предел определения	Рассчитать исходя из полученной калибровки. Можно проверить предел хроматографированием раствора с рассчитанным пределом детектирования.	<ol style="list-style-type: none">1. Рассчитать LOD2. Рассчитать LOQ3. Привести хроматограммы

Поэтапная валидация метода

Этап валидации	Как выполнить ?	Что должно быть задокументировано ?
Робасность Надежность	В несколько различных условиях (разные приборы, разные дни, разные температуры окружающей среды, разные реактивы и т.д.), проанализировать раствор одного и того же препарата.	<ol style="list-style-type: none">1. RSD пиков для разных условий.2. RT основных компонентов смесей.3. Эффективность по каждому из пиков.4. Коэффициенты асимметрии пиков.5. Разрешение для критичных пар пиков6. Привести хроматограммы
Неопределенность	Расчитать неопределенность одним из приведенных методов , согласно модельного уравнения	<ol style="list-style-type: none">1. Привести бюджет неопределенности (в случае расчета по GUM).2. Привести значение расширенной неопределенности, стандартной суммарной неопределенности, коэффициента охвата.

Поэтапная валидация метода

Итак, как же спланировать серии экспериментов для получения исчерпывающего отчета по валидации метода ?

Критерии, которые можно рассчитать из данной серии анализов	Планирование серий анализов							
Диапазон , Линейность, LOD, LOQ	<u>Образцы</u> : Растворы стандартного вещества различной концентрации.							
	Повторности							
	Т	С	1	2	3	...	9	10
	о	C _{min}	*	*	*	*	*	*
	ч	C ₂	*	*	*	-	-	-
	к	C ₃	*	*	*	-	-	-
	а	C ₄	*	*	*	-	-	-
л	C _{max}	*	*	*	*	*	*	
и	Итог							
б								
р								
о								
в								
к								
и								

Поэтапная валидация метода

Критерии, которые можно рассчитать из данной серии анализов

Точность, сходимость, промежуточная прецизионность, неопределенность, рикавери

Планирование серий анализов

Образцы: Растворы различной концентрации, полученные путем обработки модельных смесей, полученных на основе плацебо с добавкой определяемых веществ

		Первый оператор или прибор . Повторности					
То чк и ка ли бр ов ки	С, % от номинала	1	2	3	...	9	10
	C_{\min} (60%)	*	*	*	-	-	-
	C_2 (80 %)	*	*	*	-	-	-
	C_3 (100 %)	*	*	*	*	*	*
	C_4 (120 %)	*	*	*	-	-	-
	C_{\max} (140 %)	*	*	*	-	-	-

Итого : 25 обязательных инъекций для одного оператора и прибора.

		Второй оператор или прибор . Повторности					
С, % от номинала	1	2	3	...	9	10	
	C_3 (100 %)	*	*	*	*	*	*

Итого : 10 обязательных инъекций для второго оператора и прибора.



Благодарю за внимание !

