ЛЕКЦИЯ 9

ЗНАЧАЩИЕ ЦИФРЫ И ПРАВИЛА ОБРАЩЕНИЯ С НИМИ ПРИ РАСЧЕТАХ В АНАЛИТИЧЕСКОЙ ХИМИИ

ЗАКОН РАСПРОСТРАНЕНИЯ ПОГРЕШНОСТЕЙ

МЕТРОЛОГИЧЕСКИЕ АСПЕКТЫ АНАЛИЗА

СТАТИСТИЧЕСКИЕ МЕТОДЫ ПРОВЕРКИ ГИПОТЕЗ

ЗНАЧАЩИЕ ЦИФРЫ И ПРАВИЛА ОБРАЩЕНИЯ С НИМИ

- *Значащие цифры – это все достоверно известные цифры данного числа плюс первая недостоверная цифра.*
- *Достоверность результатов аналитических определений определяется реальными возможностями используемого метода или методики.*
- *В качестве статистических критериев оценки недостоверности результатов анализа служит стандартное отклонение или доверительный интервал.*
- *При отсутствии конкретных данных считают, что недостоверность последней цифры равна ± 1.*

- *Нуль, стоящий в середине или в конце числа, всегда является значащей цифрой.*

- *При записи результата анализа нули, стоящие в конце числа следует исключить, если они не являются значащими, а результат представить в виде числа с нормальной формой представления:*

$$a \cdot 10^{n}, \text{ где } 1 < a < 10.$$

Пример.

- *Если число 1000 содержит 3 значащие цифры, то правильная запись - $1{,}00 \cdot 10^3$.*
- *Если в нем 4 значащие цифры, то число следует записать как $1{,}000 \cdot 10^3$.*

- *Нули, стоящие в начале числа, не являются значащими.*

Пример. Масса образца равна 0,1 г. Если взвешивание проводили на аналитических весах с погрешностью ± 0,0002 г, то правильное представление результата – 0,1000 г или $1,000 \cdot 10^{-1}$ г.

- *Незначащие цифры исключают, округляя число.*

- *Если отбрасываемая цифра меньше 5, то последняя значащая цифра не изменяется.*
- *Если отбрасываемая цифра больше или равна 5, то последняя значащая цифра увеличивается на единицу.*
- *Если за первой недостоверной цифрой следует цифра 5, то применяют и другое правило: округляют цифру 5 до ближайшего четного числа.*

- *При расчете результатов анализа часто обращаются к атомным массам элементов. Атомные массы определены довольно точно, хотя и с погрешностью.*

- *Недостоверность последней цифры в атомных массах равна ± 1 или ± 3, если она выделена мелким шрифтом.*

- *При расчетах атомные и молярные массы должны быть взяты со всеми значащими цифрами. Округление их до целых значений приводит к грубому искажению (завышению или занижению) результата анализа.*

- *Количество значащих цифр в числе, полученном в результате вычислений, определяется из сравнения недостоверности чисел.*

- *При сложении и вычитании нескольких чисел в результате вычисления оставляют столько **цифр после запятой**, сколько их содержится в слагаемом с наименьшим числом десятичных знаков.*

- *При умножении и делении количество **значащих цифр** конечного результата определяется числом, имеющим наибольшую относительную недостоверность и наименьшее количество значащих цифр.*

- *При возведении числа в степень **относительная недостоверность** результата увеличивается в число раз, равное степени.*

- ✓ *Логарифмические величины должны содержать в мантиссе такое же количество значащих цифр, как и использованные при расчете нестепенные числа. Цифры, полученные при логарифмировании степенного члена, не являются значащими.*

- ❖ *Во избежание накопления ошибки округления в промежуточных расчетах следует сохранять на одну цифру больше, чем требуют правила обращения со значащими цифрами. В конечном результате эта цифра округляется.*

- ❖ *Правила обращения со значащими цифрами при математических операциях можно строго обосновать на основе закона распространения погрешностей.*

- *Результаты грави- и титриметрических определений в большинстве случаев записывают в виде чисел, содержащих* ***четыре значащие цифры****, что связано с погрешностью измерения массы веществ (±0,0002 г) и объемов (±0,03 мл) растворов.*

- *Результат анализа и его погрешность должны содержать одинаковое число знаков после запятой.*

- *В величине s, $\Delta_{дов}$ следует оставлять две значащие цифры, если первая цифра в них меньше трех, и одну значащую цифру, если она больше или равна трем. В последнем случае результат анализа следует округлить.*

ЗАКОН РАСПРОСТРАНЕНИЯ ПОГРЕШНОСТЕЙ

- *Аналитические определения редко бывают прямыми; чаще искомое значение содержания находят косвенно, по его зависимости от двух или большего числа величин, найденных прямыми измерениями.*

- *Каждая из стадий анализа характеризуется своей погрешностью, которая входит в общую погрешность анализа.*

Косвенную величину y можно представить в общем виде как функцию аргументов – экспериментально измеряемых величин x_1, x_2,…x_n (масса навески, объем аликвоты, объем титранта):

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

Если аргументы не зависят друг от друга, а оценки их дисперсий $s_{x_1}^2, s_{x_2}^2, ..., s_{x_n}^2$ известны, то оценка дисперсии s_y^2 может быть вычислена по закону распространения погрешностей:

$$s_y^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 \cdot s_{x_1}^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \cdot s_{x_2}^2 + ... + \left(\frac{\partial f}{\partial x_n}\right)^2 \cdot s_{x_n}^2.$$

- *Если y - функция 2 случайных экспериментально измеряемых величин x_1 и x_2, то для распространения погрешностей можно записать следующие формулы (a, b – константы, значения которых точно известны):*

- *для суммы и разности* $y = ax_1 \pm bx_2$

$$s_y^2 = a^2 s_{x_1}^2 + b^2 s_{x_2}^2,$$

- *для произведения и частного* $y = ax_1 bx_2$, $y = ax_1 / bx_2$

$$s_r^2(y) = s_r^2(x_1) + s_r^2(x_2), \quad \text{где} \;\; s_r(y) = s_y / y, \;\; s_r(x_i) = s_{x_i} / x_i\,.$$

- *Таким образом, при сложении и вычитании суммируются дисперсии абсолютных погрешностей (s^2), а при умножении и делении складываются дисперсии относительных погрешностей (s_r^2).*

•*для возведения в степень n или извлечения корня n-степени*

$$y = ax^n$$

$$s_r(y) = n\, s_r(x), \quad где \;\; s_r(y) = s_y / y, \;\; s_r(x) = s_x / x.$$

•*для логарифмирования*

$$y = a \lg x = \frac{a}{2{,}3} \ln x$$

$$s_y = \frac{a}{2{,}3} s_r(x), \quad где \;\; s_r(x) = s_x / x.$$

• *для расчета антилогарифма*

$$y = a \cdot 10^x = a \cdot e^{2{,}3x}$$

$$s_r(y) = 2{,}3 \cdot s_x, \quad где \;\; s_r(y) = s_y / y.$$

Минимально допустимая масса навески в гравиметрии

- *Массу весовой формы $m_{в.\ ф.}$ находят по разности масс тигля с осадком (m_2) и пустого тигля (m_1) :*

$$m_{в.\ ф.} = m_2 - m_1$$

- *Из закона распространения погрешностей следует, что абсолютная погрешность (s) массы весовой формы равна:*

$$s^2_{m_{в.ф.}} = s^2_{m_1} + s^2_{m_2}, \qquad s_{m_{в.ф.}} = \sqrt{s^2_{m_1} + s^2_{m_2}}.$$

- *Абсолютные погрешности массы тигля с осадком и пустого тигля определяются погрешностью взвешивания аналитических весов (± 0,1÷ 0,2 мг), тогда*

$$s_{m_{в.ф.}} = \pm\sqrt{2}\,(0{,}1 \div 0{,}2) \cong \pm\,0{,}2\ \text{мг} = 2\cdot 10^{-4}\ \text{г}.$$

- Относительная погрешность массы весовой формы (s_r) не превысит ± 0,1 %, если масса весовой формы осадка равна как минимум 0,2 г:

$$s_r = \frac{s_{m_{в.ф.}}}{m_{в.ф.}} = \frac{\pm 2 \cdot 10^{-4}}{m_{в.ф.}} \leq \pm 1 \cdot 10^{-3}, \qquad m_{в.ф.} \geq 0{,}2 \text{ г.}$$

МЕТРОЛОГИЧЕСКИЕ АСПЕКТЫ АНАЛИЗА

- *Анализ является метрологической процедурой, так как получение результата связано с измерениями.*
- *Метрология – наука об измерениях и путях достижения необходимого уровня их точности.*
- *Измерение – установление искомого значения определяемой величины с помощью специальных технических средств.*
- *Метрологические характеристики результатов анализа являются показателями качества выполненных определений.*

- *Погрешность любого измерения δ имеет 2 вклада – систематическая погрешность $\delta_{сист}$ и случайная $\delta_{сл}$:*

$$\delta = \delta_{сист} + \delta_{сл}$$

- *Систематическая погрешность, в отличие от случайной, имеет знак.*

- *Для минимизации систематической погрешности необходимо:*
  - *выбрать оптимальную методику измерений,*
  - *устранить влияние мешающих компонентов,*
  - *поверить измерительные приборы.*

- *Правильность – это степень близости среднего результата серии наблюдаемых значений к значению, принятому за опорное:* $\delta_{сист} = \bar{x} - \mu.$

Способы проверки правильности анализа:

- *способ «введено – найдено»;*
- *анализ стандартных образцов с точно известным содержанием компонента;*
- *проведение анализа разными методами;*
- *варьирование размера пробы.*

Для устранения неизвестной систематической погрешности используются следующие приемы:

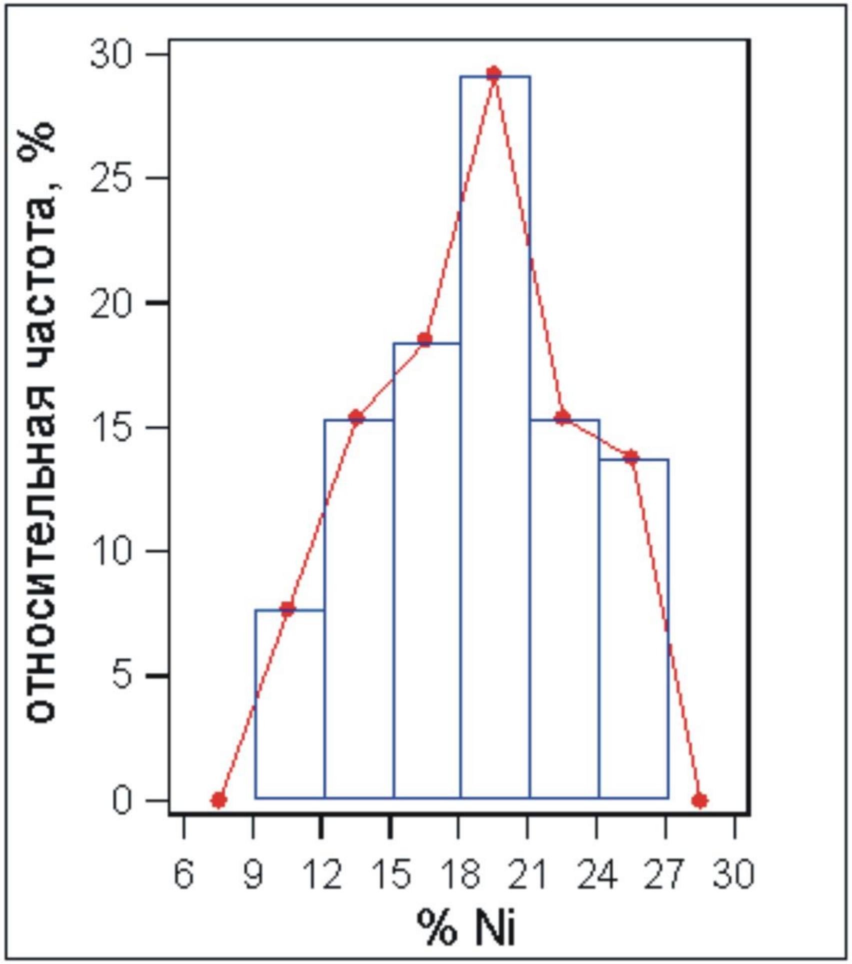
- *релятивизация – прием, при котором аналитическое определение проводят относительно другого объекта (в этом случае погрешности вычитаются;*
- *рандомизация (систематические погрешности переводят в разряд случайных, для этого, например, используют разные пипетки для отбора аликвот).*

- ✓ *Прецизионность – степень разброса независимых результатов измерений относительно среднего.*
- ✓ *Сходимость (внутрилабораторная воспроизводимость) характеризует близость результатов анализа одного и того же образца, полученных одним и тем же методом в одних и тех же условиях в течение короткого промежутка времени.*
- ✓ *Межлабораторную воспроизводимость характеризует близость результатов анализа одного и того же образца, полученных одним и тем же методом <u>в различных условиях</u> (в нескольких лабораториях, на разных приборах).*
- ✓ *Меры разброса: стандартное отклонение(s), дисперсия (s^2), относительное стандартное отклонение (s_r), размах варьирования (R_x).*

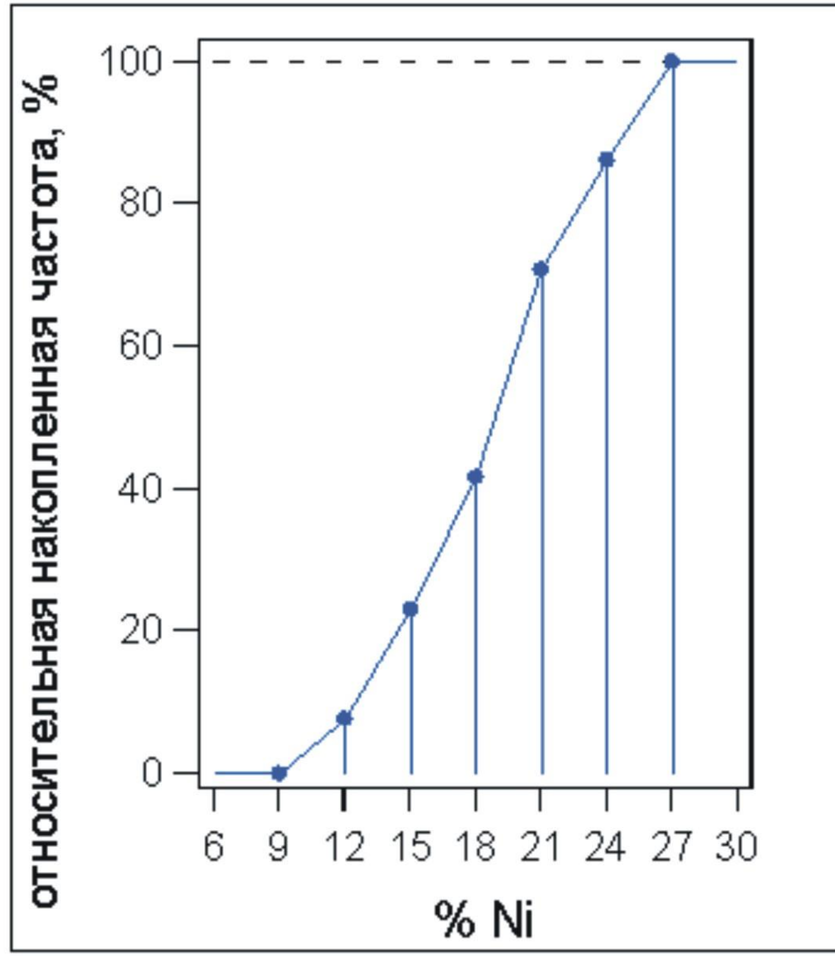
Гистограмма и график относительных частот

Пример: N=65 образцов сплава никеля

%*Ni* (≥ – <)	число образцов в интервале, n_i	относительная частота (n_i/N) в интервале, %	относительная накопленная частота (Σ n_i/N) в интервале, %
9 – 12	5	7,7	≤ 12 7,7
12 – 15	10	15,4	≤ 15 23,1
15 – 18	12	18,5	≤ 18 41,6
18 – 21	19	29,2	≤ 21 70,8
21 – 24	10	15,4	≤ 24 86,2
24 – 27	9	13,8	≤ 27 100,0



относительная частота, %
30
25
20
15
10
5
0
6
9
12
15
18
21
24
27
30
% Ni

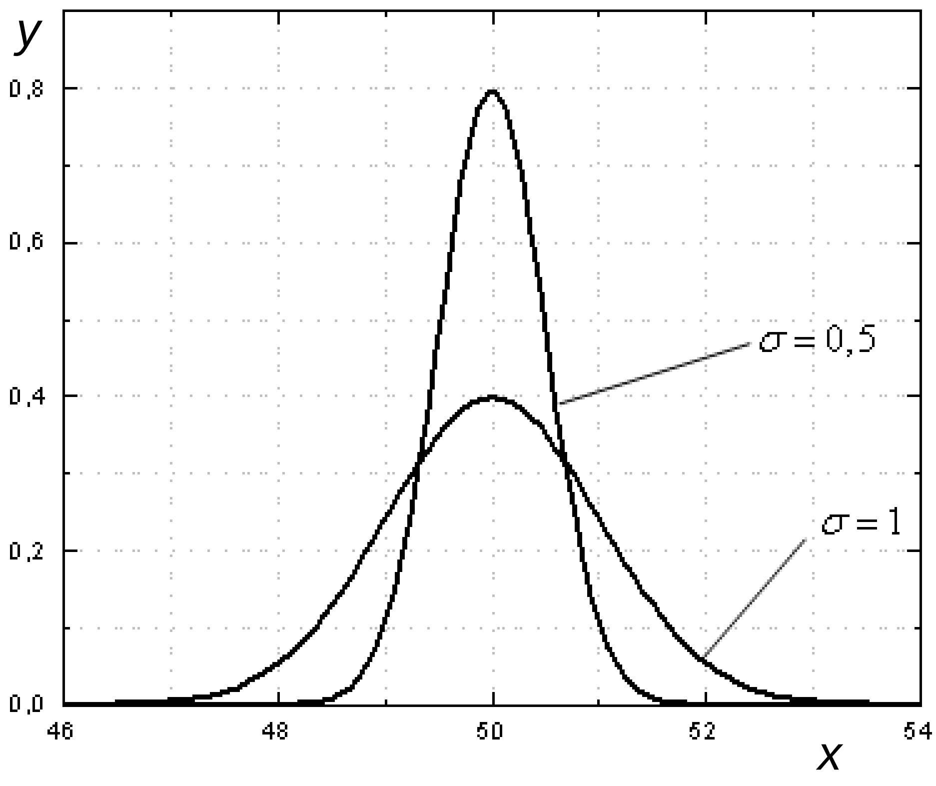


относительная накопленная частота, %
100
80
60
40
20
0
6
9
12
15
18
21
24
27
30
% Ni

Нормальный закон распределения (кривая Гаусса)

- ✓ *Плотность вероятности h появления значения x_i в серии измерений можно описать функцией Гаусса:*

$$y = h(x_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right].$$



- ✓ Вероятность попадания единичного значения x_i в интервал $\mu \pm dx$ называется <u>доверительной вероятностью</u> ($P_{дов}$) и равна площади под кривой нормального распределения, рассчитанной в этих пределах:

$$P_{дов} = \int_{\mu-dx}^{\mu+dx} h(x_i)\, dx.$$

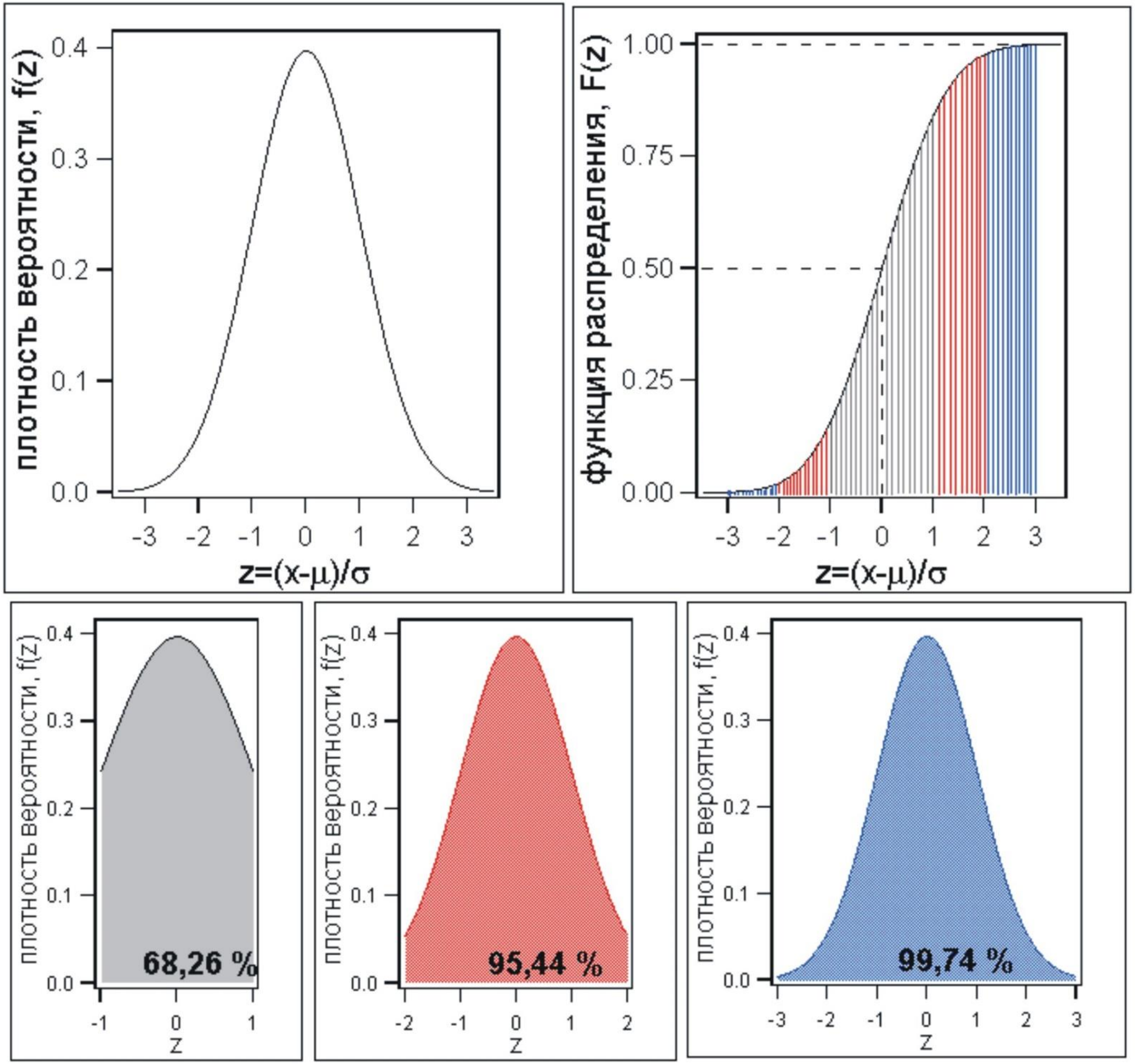
- ✓ Доверительная вероятность характеризует вероятность того, что ожидаемая величина x_i лежит внутри некоторого интервала, называемого <u>доверительным</u>.

- ✓ Для большинства аналитических определений принимают $P_{дов} = 0{,}95$.

- ✓ *Вероятность того, что единичное значение x_i не попадет в интервал μ ± dx, называют уровнем значимости* α *:* $\alpha = 1 - P_{дов}$
- ✓ *Для большинства аналитических определений уровень значимости* α *равен 0,05.*
- ✓ *Нормальное распределение часто записывают через безразмерную переменную z:*

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right], \ где \ z = \frac{x_i - \mu}{\sigma};$$

$$P_{дов} = \int_{-dz}^{dz} f(z)\,dz.$$


плотность вероятности, f(z)
z=(x-μ)/σ
функция распределения, F(z)
68,26 %
95,44 %
99,74 %
z

- ✓ *Результат анализа представляют в виде среднего значения с доверительным интервалом* $\Delta_{дов}$ *:*

$$\mu = \bar{x} \pm \Delta_{дов}$$

- ✓ *Доверительный интервал оценивает погрешность определения среднего значения и показывает, в какой окрестности около экспериментального среднего значения может находиться истинное значение* μ*.*

- ✓ *Если число наблюдений велико (n>30), то величина доверительного интервала рассчитывается по формуле:*

$$\Delta_{дов} = \pm 1{,}96\, \sigma / \sqrt{n} \quad (P_{дов} = 0{,}95);$$

$$\Delta_{дов} = \pm 2{,}58\, \sigma / \sqrt{n} \quad (P_{дов} = 0{,}99).$$

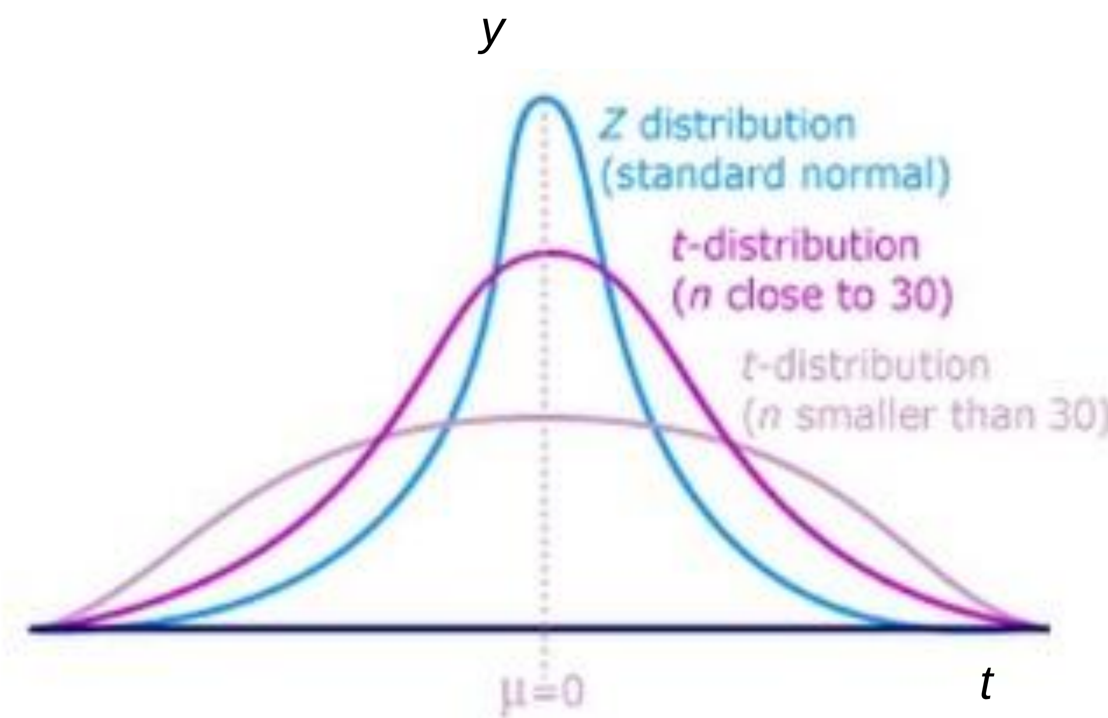
Распределение Стьюдента

- *Для выборочной совокупности с небольшим числом определений значения μ и σ неизвестны.*
- *Вместо этих величин используют их <u>оценки</u> :* $\bar{x}, s.$
- *Для малых выборок (n < 30) применяется специально построенное t-распределение (распределение Стьюдента, 1908 г, В. Госсет).*
- *Величина t является мерой отклонения среднего результата измеряемой величины от ее математического ожидания в единицах выборочного стандартного отклонения:*

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{(\bar{x} - \mu)\sqrt{n}}{s}.$$

Распределение Стьюдента связывает значение переменной t со значением плотности вероятности y появления этого значения в серии измерений

$$y = f(t,n) = B_{n-1}\left(\frac{1+t^2}{n-1}\right)^{-n/2}$$



- *Результат анализа представляют в виде среднего значения с доверительным интервалом $\Delta_{дов}$:*

$$\mu = \bar{x} \pm \Delta_{дов}$$

- *Величина доверительного интервала для небольшой выборочной совокупности (n < 30) рассчитывается по формуле:*

$$\Delta_{дов} = \pm t(P_{дов}, f)\, s/\sqrt{n},\ где\ \ f = n-1.$$

- *Необходимое число параллельных определений:*

f	*t(0,95; f)*	$\Delta_{дов}$	*f*	*t(0,95; f)*	$\Delta_{дов}$
1	*12,7*	*8,98s*	*5*	*2,57*	*1,05s*
2	*4,30*	*2,48s*	*6*	*2,45*	*0,93s*
3	*3,18*	*1,59s*	*7*	*2,37*	*0,84s*
4	*2,78*	*1,24s*	*8*	*2,31*	*0,77s*

n = 6 - 9

СТАТИСТИЧЕСКИЕ МЕТОДЫ ПРОВЕРКИ ГИПОТЕЗ

❖ *Выявление грубых ошибок*

✓ *Промахи – результаты, полученные с нарушением методики или являющиеся следствием грубых ошибок, допущенных в ходе анализа (потери осадка, неправильный отбор аликвот и т. п.).*

✓ *Если выборка содержит выпадающий результат, причина появления которого неизвестна, то его называют выбросом.*

✓ *Для исключения выбросов используют:*

- *если $n > 10$, то 3s-критерий,*
- *если $3 < n < 10$, то Q-критерий.*

✓ *Из выборки можно исключить не более 1/3 результатов.*

❖ *Сравнение среднего значения и константы (простой t-тест Стьюдента)*

✓ *<u>Нуль-гипотеза</u>: экспериментальное среднее значение и значение, принятое за опорное, представляют собой одно и то же значение, различие между ними носит случайный характер.*

✓ *Альтернативная гипотеза: различие между средним значением и опорным является следствием действия систематических погрешностей, искажающих результат анализа.*

✓ *Вероятность ошибочного отклонению нуль-гипотезы (или ошибочного принятия альтернативной) в анализе обычно принимают равной 0,05.*

✓ *Нуль-гипотеза принимается, если*

$$\left|\bar{x} - \mu\right| \leq \Delta_{дов} = t(P_{дов}, f) \cdot s / \sqrt{n}$$

- *Сравнение двух экспериментальных средних*
- *1 этап – сравнение выборочных дисперсий*
- *Нуль-гипотеза: выборочные дисперсии являются оценками дисперсии одной и той же генеральной совокупности, различие между ними носит случайный характер. Для сравнения используют F-распределение Фишера*.
- *Нуль-гипотеза принимается, если*

$$F_{эксп} \le F_{табл}(P_{дов}, f_1, f_2), \text{ где } F_{эксп} = s_1^2 \big/ s_2^2 \qquad (s_1^2 > s_2^2).$$

- *В таблице значений критерия Фишера число степеней свободы f_1 (выборки с бóльшей дисперсией) приведено в горизонтальном ряду, f_2 – в вертикальном ряду.*

✓ 2 этап – сравнение выборочных средних (расширенный t-тест Стьюдента)

- Нуль-гипотеза: выборочные средние значения являются оценками среднего значения одной и той же генеральной совокупности, различие между ними носит случайный характер.

- Нуль-гипотеза принимается, если

$$\bar{s}^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2};$$

$$\left|\bar{x}_1 - \bar{x}_2\right| \le \Delta_{дов} = t(P_{дов}, f) \cdot \bar{s} \cdot \sqrt{\frac{n_1 + n_2}{n_1\, n_2}}, \quad где \quad f = n_1 + n_2 - 2.$$

- Результаты выборок объединяют, рассчитывают среднее значение, стандартное отклонение, доверительный интервал ($f = n_1 + n_2 - 1$).