

# Part III-B: Analysis Chemistry

Lecture by 王敏

Note by THF

2024 年 9 月 24 日

## 目录

1	概论	1
2	误差和分析数据处理	2
2.1	准确度和精密度	3
2.1.1	准确度和误差	3
2.1.2	精密度和偏差	4
2.1.3	误差	5
2.2	如何提高分析结果的准确度	6
2.3	有效数字	7
2.3.1	修约规则	7
2.3.2	运算规则	8
2.4	有限量分析数据的统计处理	8
2.4.1	偶然误差的正态分布	8
2.4.2	T 分布	9
2.4.3	平均值的精密度和置信区间	10

## 1 概论

20 世纪 20-30 年代：分析化学出现四大反应平衡理论的建立

20 世纪 40-50 年代: 光电色谱仪器设备出现

**Notation.** Bloch F and Purcell E M 建立核磁共振测定

Martin A J P and Synge R L M 建立气相色谱

Heyrovsky J 建立极谱分析法

20 世纪 70 年代以来: 计算机参与自动化

**Notation.** 分析化学分析方法: 3S+2A

3S: Sensitivity, Selectivity, Speediness

2A: Accuracy, Automatics

分析化学主要发展趋势:

{  
在线分析  
原位分析  
实时分析  
活体分析  
...

**Notation.** 分析过程的步骤:

1. 分析方法选择
2. 取样 (Sampling, 具代表性的样本)
3. 制备试样 (适合与选定的分析方法, 消除可能的干扰)
4. 分析测定 (优化条件, 仪器校正, 方法验证)
5. 结果处理和表达 (统计学分析, 测量结果的可靠性分析, 书面报告)

**Notation.** 制备试样首先需要进行样品前处理

方法验证: 线性性, 灵敏性等

## 2 误差和分析数据处理

**Notation.** “挑数据”: 做标准曲线

标品浓度 (0,1,2,3,4,5,6,7)

吸光度 (0,0.1,0.2,0.3,0.4,0.001,0.6,0.7)

对于可疑数据, 需要通过其他方法进行确认 (至少 3 次试验)

## 2.1 准确度和精密度

### 2.1.1 准确度和误差

**Definition.** 准确度 (Accuracy): 测量值和真值的接近程度, 准确度的高低用误差大小衡量

**Definition.** 误差 (error): 测量结果和真值的差值

误差具有客观性和普遍性

实验结果都有误差, 测量值只能尽可能接近真实值

**Definition.** 约定真值: 由国际计量大会定义的单位及我国的法定计量单位

**Example.** 国际单位制基本单位“米”、“克”等

**Notation.** 约定真值是有一个量的真值的近似值, 误差可以忽略不计

**Definition.** 标准值: 采用可靠的分析方法、在不同实验室、由不同的分析人员、对同一个试样反复多次测定后将大量数据用数理统计求得的测量值

误差的表示方法:

$$\delta = x - \mu.$$

**Definition.** 绝对误差:  $\delta$

测量值:  $x$

真值:  $\mu$

相对误差: RE%

绝对误差 (Absolute Error) 可正可负, 单位为测量值的单位

绝对误差的绝对值越小, 准确度越高

相对误差 (Relative Error, RE):

$$\text{RE}\% = \frac{\delta}{\mu} \times 100\%.$$

或:

$$\text{RE}\% = \frac{\delta}{x} \times 100\%.$$

相对误差无单位, 可正可负

**Example.** 有真实值为 0.0020g 和 0.5000g 的两个样品, 称量结果分别为 0.0021g 和 0.5001g, 计算相对误差和绝对误差

$$\delta_1 = 0.0021 - 0.0020 = 0.0001, \delta_2 = 0.5001 - 0.5000 = 0.0001 = \delta_1$$

$$\text{RE}_1\% = \frac{0.0001}{0.0020} \times 100\% = 5\%$$

$$\text{RE}_2\% = \frac{0.0001}{0.5000} \times 100\% = 0.02\%.$$

**Notation.** RE 要求: 测高含量组分, RE 可小; 测低含量组分, RE 可大  
高含量组分对应化学分析法; 低含量组分对应仪器分析法

### 2.1.2 精密度和偏差

**Definition.** 精密度: 在规定的测定条件下, 多次平行测定结果相互吻合的程度, 精密度高低用偏差衡量

偏差: 单个测量值和测量平均值的差距

**Definition.** 绝对偏差:  $d$

相对偏差:  $d\%$

平均偏差:  $\bar{d}$

标准偏差: SD

相对标准偏差: RSD

绝对偏差:

$$d = x_i - \bar{x}.$$

相对偏差:

$$d\% = \frac{d}{\bar{x}}.$$

平均偏差:

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n}.$$

相对平均偏差:

$$\frac{\bar{d}}{\bar{x}} \times 100\% = \frac{\sum |x_i - \bar{x}|}{n \cdot \bar{x}} \times 100\%.$$

标准偏差:

$$S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (n \leq 20).$$

$n-1$  称为自由度

相对标准偏差 (RSD, 变异系数):

$$\text{RSD} = \frac{S_x}{\bar{x}} \times 100\%.$$

RSD 越小, 数据越集中, 精密度越高

**Notation.** 方法的精密度考察:

1. 重复性 (repeatability): 同一实验室, 较短时间间隔, 同一分析人员对同一试样测定所得结果的接近程度
2. 中间精密度 (intermediate precision): 改变某些试验条件对同一试样测定结果的接近程度
3. 重现性 (reproducibility): 不同实验室, 不同人员对同一试样测试结果的接近程度

**Notation.** 精密度和准确度的关系:

1. 精密度高是准确度高的前提
2. 精密度高, 准确度不一定高
3. 只有精密度和准确度都高的数据才可取

**Notation.** 准确度体现结果的正确性, 精密度体现结果的重现性

### 2.1.3 误差

**Notation.** 误差分类: 系统误差、偶然误差 (随机误差)

**Definition.** 系统误差 (可定误差): 分析中由某些确定原因造成的误差

特点:

1. 重现性
2. 单向性
3. 可以校正消除
4. 影响准确度

**Notation.** 系统误差分类:

$$\left\{ \begin{array}{l} \text{方法误差: 方法不完善} \\ \text{仪器误差: 仪器本身缺陷} \\ \text{试剂误差: 试剂有杂质} \\ \text{操作误差、主观误差} \end{array} \right.$$

**Definition.** 偶然误差 (随机误差): 由偶然因素影响

特点:

1. 无单向性 (方向大小不确定)
2. 符合统计学规律 (大误差出现的概率小, 随机误差的正态分布)
3. 可以通过增加平行测定次数减小
4. 影响精密度

**Definition.** 过失: 由分析工作者粗心大意造成, 可以避免

## 2.2 如何提高分析结果的准确度

1. 选择合适的分析方法:

**Notation.** 对于化学分析法: 适用于常量组分 ( $>1\%$ ) 的分析

对于仪器分析法: 适用于微量组分 ( $0.01 \sim 1\%$ ) 或痕量组分 ( $<0.01\%$ ) 的分析

2. 减少相对误差: 增大取样量
3. 减少偶然误差: 增加平行测试次数 ( $\geq 3$ , 活体测试  $\geq 6$ )
4. 消除系统误差:

**Notation.** 常用方法:

1. 与经典方法比较: (用其他方法测试该方法)
2. 对照试验 (control test): (用其他样本对比)
3. 回收试验/标准加入法: 测量原样本  $x_1$ , 加入某标准量  $x_2$ , 测量加入后的样本  $x_3$ , 计算回收率:

$$\text{Recovery}\% = \frac{x_3 - x_1}{x_2} \times 100\% \approx 95\% \sim 105\%.$$

4. 空白试验: 不加试样测试, 得到并扣去空白值, 用于检验由试剂、容器等引入杂质导致的系统误差

5. 仪器校正

## 2.3 有效数字

**Definition.** 有效数字: 分析工作中实际上能测量到的数字

原则上:

1. 在记录测量数据时, 只允许保留一位可疑数字 (欠准数字)
2. 误差是末尾数的  $\pm 1$  个单位

**Example.** 在分析天平上称出  $m = 21.5370\text{g}$ , 则真值为  $21.5370 \pm 0.0001\text{g}$

**Notation.** 有效数字的末尾 0 不可省略: 反映相对误差

**Question.** 如何判断有效数字的位数?

1. 在数据中 1 ~ 9 均为有效数字 (0 待定)
2. 算式中的倍数、分数及某些常数 ( $\pi, e$  等) 可看为无数位有效数字
3. 变化单位时有效数字的位数必须保持不变, 如  $0.0015\text{g} = 1.5\text{mg}$
4. pH 和  $\text{pK}_a$  等对数值, 有效数字仅取决于小数部分, 例:  $\text{pH} = 4.23$ , 有效数字 2 位

**Notation.** 0 的位置和有效数字:

1. 小数前的 0 起定位作用, 后面的 0 为有效数字
2. 整数后的 0 不一定是有效数字

**Example.**  $36000 \Rightarrow 3.60 \times 10^4$ : 3 位有效数字

### 2.3.1 修约规则

1. 四舍六入五留双, 五后有数需进位
2. 修约标准偏差: 只进不舍, 降低精密度, 提高可信度
  - 2.1. 可多保留一位有效数字进行计算
  - 2.2. 与标准限度值比较不应修约

表 1: 修约为两位

原值	修约值	原因
3.249	3.2	四舍
8.361	8.4	六入
6.550	6.6	五留双
6.250	6.2	五留双
6.252	6.3	五后有数

### 2.3.2 运算规则

1. 加减法: 结果的小数位数以小数点后位数最少的为标准

**Example.**

$$0.0121 + 25.64 + 1.057 = 26.7091 \approx 26.71.$$

2. 乘法: 取相对误差最大的为标准

**Example.**

$$\frac{0.0325 \times 5.10 \times 60.1}{139.8} \approx 0.0712.$$

**Notation.** 对于高含量组分 ( $w > 10\%$ ), 分析结果一般保留 4 位有效数字

对于中等含量组分 ( $1\% < w < 10\%$ ), 保留 3 位

对于微量组分 ( $w < 1\%$ ), 保留 2 位

## 2.4 有限量分析数据的统计处理

### 2.4.1 偶然误差的正态分布

正态分布的概率密度函数:

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}.$$

**Notation.**  $x$ : 测量值

$\mu$ : 无限次测量的总体平均值

$\sigma$ : 总体偏差



特点:

1. 当  $x = \mu$  时,  $y$  最大: 大部分测量值集中在算术平均值附近
2. 函数图像以  $x = \mu$  的直线对称: 正负误差出现的概率相同
3.  $x \rightarrow -\infty$  或  $x \rightarrow +\infty$  时: 无限趋近  $x$  轴
4.  $\mu$  越大, 数据越分散, 函数图像矮小、坡度较缓
- 5.

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

**Notation.** 若令

$$u = \frac{x - \mu}{\sigma}.$$

则:

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

称为标准正态分布曲线

**Notation.**  $3\sigma$  标准: 测量结果需要有 99.7% 以上的数据在真值范围内, 即:

$$\int_{\mu-3\sigma}^{\mu+3\sigma} f(x) dx \approx 99.7\%.$$

## 2.4.2 T 分布

**Definition.** T 分布为有限量数据  $n$  平均值的概率误差分布

$$t = \frac{|\bar{x} - \mu|}{S} \sqrt{n}.$$

$x$ : 样本平均值

$S$ : 样品标准偏差

$t$  值随自由度  $f (f = n - 1)$  而变

当  $f \rightarrow +\infty$  时为正态分布

**Notation.** T 分布和正态分布的异同:

相同: 形状相似, 积分面积表示概率

不同: T 不同时概率不同

**Notation.** T 分布相关概念:

1. 自由度  $f$
2. 置信区间: 以测定区间为中心, 包括总体平均值在内的可信范围

$$\bar{x} \pm \frac{tS}{\sqrt{n}}.$$

3. 置信水平 (置信度  $P$ ): 样本平均值落在置信区间的概率

$$P = \mu \pm \frac{tS}{\sqrt{n}}.$$

4. 显著性水平  $\alpha = 1 - P$

$t$  值的表达: 一定  $P$  下,  $t \rightarrow t_{\alpha, f}$

### 2.4.3 平均值的精密度和置信区间

**Example.** 有一个样品,  $m$  个人各测量  $n$  次, 计算出每个人测得的平均值, 平均值的分布符合正态分布

平均值的标准偏差:

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}}.$$

即: 增加平行测定次数  $n$  可以减小平均值的标准偏差  $S_{\bar{x}}$

**Notation.** 增加测定次数, 平均值的标准偏差呈反比变化, 一般 3 到 4 次已经达到目标, 继续增加效果不显著