实验数据的误差分析

§1 实验数据的误差分析

误差的基本概念

实验数据的精准度

实验数据的真值与平均值

误差的表示法

§2 有效数的确定

实验数据的有效数与记数法

有效数的运算

化工实验数据处理

实验数据的误差分析

误差分析在化工实验研究中的重要性

通过实验测量所得大批数据是实验的首要任务

- 但在实验中，由于测量仪表和人的观察等方面的原因，实验数据总存在一些误差，所以在整理数据时，首先应对实验数据的可靠性进行客观的评定。
- 误差分析的目的就是评定实验数据的精确性或误差，通过误差分析，可以认清误差的来源及其影响，并设法排除数据中所包含的无效成分，还可进一步改进实验方案。在实验中注意哪些是影响实验精确度的主要方面，细心操作，从而提高实验的精确性。

误差的基本概念

实验数据的误差来源及分类

误差是实验测量值（包括间接测量值）与真值（客观存在的准确值）之差别，基于下列原因，误差可分为三类：

➢ **系统误差**

由于测量仪器不良，如刻度不准，零点未校准；或测量环境不标准，如温度、压力、风速等偏离校准值；实验人员的习惯和偏向等因素所引起的系统误差。

这类误差在一系列测量中，大小和符号不变或有固定的规律，经过精确的校正可以消除。

➢ 随机误差（偶然误差）

是由一些不易控制的因素所引起的，如测量值的波动，肉眼观察欠准确等。这类误差在一系列测量中的数值和符号是不确定的，而且是无法消除的，但它服从统计规律，也是可以认识的。

➢过失误差

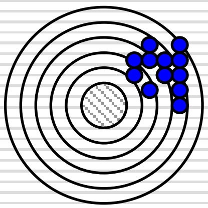
主要是由实验人员粗心大意，如读数错误、记录错误或操作失误所致。这类误差往往与正常值相差很大，应在整理数据时加以剔除。

实验数据的精准度

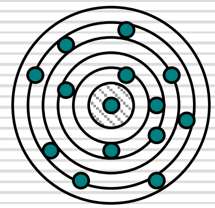
精准度与误差是相反的，精确度高，误差就小；误差大，精确度就低。精确度表示测量结果与其值接近程度。

精密度：测量中所得到的数据重复性的大小。

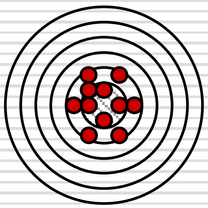
- 它反应随机误差的大小，以打靶为例



弹着点密集而离靶心（真值）甚远，说明精密度高，随机误差小，但系统误差大



随机误差大，但系统误差较小，即精密度低而正确度较高



系统误差与随机误差均小，精确度高。精确度高则精密度与正确度均高。

实验数据的真值与平均值

- 真值是待测物理量客观存在的确定值，由于测量时不可避免地存在一定误差，故真值是无法测得的。
- 但是在消除系统误差后，经过足够数次测定，根据随机误差中正负误差出现几率相等的规律，测定结果的平均值可以无限接近真值。
- 而实际上测量次数总是有限的，由此得出的平均值只能近似于真值，称此平均值为最佳值。
- 计算中可将此最佳值当作真值，或用“标准仪表”（即精确度较高的仪表）所测之值当作真值。
- 化工中常用的平均值有4种：

算术平均值 x_m

设$x_1,x_2,x_3,\ldots x_n$ 为各次测量值, n为测量次数，则算术平均值为：

$$x_m = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

算术平均值是最常用的一种平均值，因为测定值的误差分布一般服从正态分布，可以证明算术平均值即为一组等精度测量的最佳值或最可信赖值。

几何平均值 x_c

$$x_c = \sqrt[n]{x_1 x_2 \dots x_n}$$

均方根平均值 x_s

$$x_s = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$

对数平均值 x_l

$$x_l = \frac{x_1 - x_2}{\ln \dfrac{x_1}{x_2}}$$

对数平均值多用于热量和质量传递中，当 $x_1/x_2 < 2$

时，可用算术平均值代替对数平均值，引起的误差 <4%

误差的表示法

绝对误差 d　　　残余误差

某物理量在一系列测量中，某测量值与其真值之差称绝对误差。

实际工作中常以最佳值代替真值，则

$$d_i = x_i - X \approx x_i - x_m$$

式中：

d_i——绝对误差；

x_i——i 次测量值；

X——真值；

x_m——平均值。

如在实验中对物理量的测量只进行一次，可根据测量仪器出厂鉴定书注明的误差，或可取仪器最小刻度值的一半作为测量的误差。

□例如某压力表注明精（确）度为1.5级，即表明该仪表最大误差为相当档次最大量程之1.5%，若最大量程为0.4MPa，该压力表最大误差为：

$$0.4\times1.5\%\,MPa=0.006\,MPa=6\,kPa=6\times10^{3}\,Pa$$

□某天平的感量或名义分度值为0.1mg，则表明该天平的最小刻度或有把握正确读得的最小单位为0.1mg，即最大误差为0.1mg。

□化工原理实验中最常用的U形管压差计、转子流量计、秒表、量筒、电压表等仪表原则上均取其最小刻度值为最大误差，而取其最小刻度值的一半作为绝对误差。

相对误差 ε

为了比较不同测量值的精确度，以绝对误差与真值（或平均值）之比作为相对误差：

$$\varepsilon = \frac{d}{|X|} \approx \frac{d}{x_m} \times 100\%$$

在单次测量中

$$\varepsilon = \frac{d}{x_i} \times 100\%$$

式中：d——绝对误差；

$|X|$——真值的平均值；

x_m——平均值。

例：压力测量

今欲测量大约8kPa（表压）的空气压力，实验所用仪表有：

（1）1.5级，量程 0.2MPa的弹簧管式压力表；

（2）标尺分度为1mm 的U形管水银柱压差计；

（3）标尺分度为1mm 的U形管水柱压差计。

求各仪表的相对误差。

(1) 压力表

绝对误差：$d = 0.2 \times 1.5\% MPa = 0.003 MPa = 3 kPa$

相对误差：$\varepsilon = 3/8 \times 100\% = 37.5\%$

(2) 水银压差计

绝对误差：$d = 0.5 \times 1 \times (101325/760)Pa = 66.65Pa$

相对误差：$$\varepsilon = \frac{66.65}{8000} \times 100\% = 0.83\%$$

(3) 水柱压差计

绝对误差：$d = 0.5 \times 1 \times (101325/10330)Pa = 4.9Pa$

相对误差：$$\varepsilon = \frac{4.9 \times 10^{-3}}{8} \times 100\% = 0.061\%$$

可见用量程较大的仪表，测量数值较小的物理量时，相对误差较大。故而，选择测试仪表十分重要。

算术平均误差 δ

是一系列测量值的误差的绝对值的算术平均值。
是表示一系列测定值误差的较好方法之一。

$$\delta = \frac{\sum |x_i - x_m|}{n} = \frac{\sum |d_i|}{n}$$

式中： x_i—测量值，i=1，2，3. . . , n;
x_m—平均值；
d_i—绝对误差

标准误差（均方误差）σ

在有限次测量中，标准误差可用下式表示：

$$\sigma = \sqrt{\frac{\sum (x_i - x_m)^2}{n-1}} = \sqrt{\frac{\sum d_i^2}{n-1}}$$

标准误差是目前最常用的一种表示精确度的方法，它不但与一系列测量值中的每个数据有关，而且对其中较大的误差或较小的误差敏感性很强，能较好地反映实验数据的精确度。实验越精确，其标准误差越小。

实验数据的有效数与记数法

有效数字

实验数据或根据测量值的计算结果，取几位数作为有效数才是有意义的呢？

这要由测量仪表的精确度而定，一般应记录到仪表最小刻度的十分之一位。

- 某液位计标尺的最小分度为1mm，则读数可以到0.1mm。
- 如在测定时液位高在刻度524mm与525mm的中间，则应记液面高为524.5mm，其中前三位是直接读出的，是准确的，最后一位是估计的，是欠准或可疑的，该数据为4位有效数。
- 如液位恰在524mm刻度上，则数据应记作524.0mm，若记作524mm，则失去了一位精确度。

总之，有效数中应有而且只能有一位（末位）欠准数字。

有效数与误差的关系：由上可见，液位高度524.5mm中，最大误差为±0.5mm，也就是说误差为末位的一半。

科学记数法

在科学与工程中，为了清楚地表示有效数或数据的精度，通常将有效数写出并在第1位数后加小数点，而数值的数量级由10的整数幂来确定，这种以10的整数幂来记数的方法称科学记数法。

例如：$0.0088=8.8\times10^{-3}$，88000（有效数3位）$=8.80\times10^{4}$

应注意，科学记数法中，在10的整数幂之前的数字应全部为有效数，位数一目了然。

例

数	有效数字位数
0.0045	2
0.004500	4
6.800×10^3	4
6.8×10^3	2
1.000	4
6800	可能是2,3,4（取决于最后面的零是否用于定位）

有效数的运算

加、减法运算

不同有效位数的数相加减，其和或差的有效数等于其中位数最少的一个，例如测得料液进出口的温度分别为65.58℃与30.4℃则

温度和：65.58℃+30.4℃=95.98℃，

温度差：65.58℃-30.4℃ =35.18℃。

结果中有**两位欠准值**，这与有效值规则不符，故第二位欠准数应舍去，按四舍五入法，其结果应为96.0℃与35.2℃。

乘、除法运算

乘积或商的有效数，其位数与各乘、除数中有效数位数最少的相同，如测得管径$D = 50.8\text{mm}$，其面积A为

$$A=\frac{\pi}{4}D^2=\frac{3.14}{4}\times 50.8^2 mm^2=2.03\times 10^3 mm^2$$

注意：π、e、g 等常数有效位数可多可少，根据需要选取。

乘方与开方计算

乘方、开方后的有效数与其底数相同。

对数计算

对数的有效数位数与其真数相同。

例如 $\lg 2.34 = 3.69 \times 10^{-1}$ $\lg 8.0 = 9.0 \times 10^{-1}$

在四个数以上的平均值计算中，平均值的有效数字可较各数据中最小有效位数多一位。

所有取自手册上的数据，其有效数按计算需要选取，但原始数据如有限制，则应服从原始数据。

一般在工程计算中取三位有效数已足够准确，在科学研究中根据需要和仪器的可能，可以取到四位有效数字。

- 从有效数的运算规则可以看到，实验结果的精确度同时受几个仪表的影响时，则测试中要使几个仪表的精确度一致，采用一两个精度特别高的仪表无助于整个实验结果精度的提高。
- 如过滤实验中，计量滤液体积的量筒分度为0.1L，而用分度为0.001秒的电子秒表时，测得27.5635s 中流过滤液1.35L，计算每升滤液通过所需要的时间为：

$$t = 27.5635s/1.35L = 27.6s/1.35L = 20.4s/L$$

可见用一个0. 1秒分度的机械秒表精度就足够了。