

偏差、误差和不确定度的准确使用

王宝金

(《原子能科学技术》编辑部 102413)

科学实验中常需对实验的结果进行数据分析, 数据分析的准确与否直接关系到实验所得到的结论是否正确, 因此数据处理对科学实验至关重要。在数据分析过程中, 常用到偏差、误差和不确定度的概念, 它们之间既相互联系, 又有区别, 因此偏差、误差和不确定度的使用极易混淆。错误使用偏差、误差和不确定度, 不仅影响到数据分析的结果, 甚至影响到由科学实验所得到的结论, 因此本文对偏差、误差和不确定度的异同进行分析, 并探讨其在数据分析时的准确使用, 以期提高数据处理的水平。

1 偏差、误差和不确定度的定义

1.1 偏差和误差

偏差为两个结果间(如标准物理常数的不同次测量结果之间, 测量结果与预期值之间)的差别, 即任意两个相关结果进行比较。误差表示观测值(或计算值)与真实值之间的差别。由偏差的定义可见, 当被比较的结果为真实值时, 偏差即为误差, 因此, 误差可认为是一种特殊偏差。由于观测值存在不确定性, 真实值一般不可知, 因此在科技论文中除非已知被比较量为真实值, 否则应慎用误差表征结果间的差别。

1.2 不确定度

不确定度用以表示测量结果中合理赋予被测量值的一个分散性参数, 即是对误差的估计。不确定度分为 A 类和 B 类: A 类是用统计方法评定的不确定度; B 类是用非统计方法评定的不确定度。A 类不确定度

一般采用贝塞尔法计算, 计算公式为:

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

式中, \bar{x} 为 n 次测量结果的算术平均值。

对实验数据进行处理时, 一般使用 A 类不确定度的计算方法, 本文仅针对 A 类不确定度。

2 偏差、误差和不确定度的区分及正确使用

对数据进行分析处理时, 由于采用偏差、误差和不确定度进行数据分析时其绝对值不能直观、明确地反映出差异的大小, 一般采用相对值, 即使用相对偏差、相对误差和相对不确定度进行数据分析。本文结合实例分析偏差、误差和不确定度的区别, 并给出简易的判别方法。

2.1 误差和偏差的区分及正确使用

由误差和偏差的定义可知, 二者的计算方法相似, 区别仅在于被比较的量是否为真实值, 除非明确真实值已知或把某一值作为真实值, 否则论文中出现的误差都应为偏差。

由误差和偏差的计算方法可知, 二者皆可能为正值或负值, 其正负表示偏离被比较量的方向, 即较被比较量是大还是小, 因此不能用误差和偏差的计算数值表征其大小, 应用其绝对值表征其大小。

下面以科技论文原文中出现的一个实例分析误差与偏差, 原文如下: “……, 从表 1 可看出, 各参数的误差均很小, 最大仅 0.11%, 可见本文计算结果得到的压力损失系数是可接受的。”

表1 不同数值对SG稳态运行参数计算结果的影响

| 参数 | 不同压力损失系数下的数值 | | 误差 /% |
|---------------------|--------------|-------|--------|
| | 75 | 86 | |
| 循环倍率 | 3.49 | 3.46 | -0.859 |
| 二次侧水质量, t | 46.88 | 46.81 | -0.149 |
| 汽体积, m ³ | 99.39 | 99.5 | 0.11 |

……”。原文中存在诸多问题。由表1可见，误差是两个计算结果的比较，其中没有也没有定义真实值，因此表1中的误差应为偏差；根据偏差的计算方法对表1的第一行参数循环倍率进行核算，经计算循环倍率的偏差为-0.03，而表1中循环倍率的误差为-0.859%，其计算方法为 $(3.46-3.49)/3.49 \times 100\% = -0.859\%$ ，因此表1中的误差应为相对偏差；原文认为相对偏差最大为0.11%，对表1中的相对偏差进行检验，取绝对值后发现最大为0.859%，因此原文描述不当，虽不影响结论，但也有小的纰漏。

2.2 误差和不确定度的区分及正确使用

不确定度是对误差的估计，二者相关，因此，本文仅分析误差和不确定度的区别。由误差和不确定度的计算方法可知：1) 误差是两个结果间的比较，不确定度是对同一实验的多次测量结果进行统计分析；

2) 误差的计算结果可能是正值或负值，而不确定度只能是正值；3) 误差的计算较简单，不确定度的计算较复杂，需要有一组数据进行计算。本文以表2为例，说明误差和不确定度间的区别。表2中是以厚度的平均值作为其真实值。由表2可明显看出误差和不确定度的区别。

表2 薄膜厚度的测量结果

| 测量次数 | 厚度 / μm | 相对误差 /% |
|--------|--------------------|---------|
| 1 | 1.55 | 1.9 |
| 2 | 1.48 | -2.6 |
| 3 | 1.56 | 2.6 |
| 4 | 1.50 | -1.3 |
| 5 | 1.53 | 0.7 |
| 平均值 | 1.52 μm | |
| 相对不确定度 | 3.4% | |

3 结论

在进行数据分析时不能正确使用偏差、误差和不确定度，给科学实验的数据分析带来困难，甚至影响到分析结果的准确程度。本文探讨了偏差、误差和不确定度的正确使用，可为科研工作者和学生在对科学实验进行数据分析时提供一定的参考。



科苑快讯

低成本钻石

碳的一种新型同素异形体，被称为“Q碳”(Q-carbon)，硬度甚至高于钻石。这种新的物质形式是由美国罗利市北卡罗来纳州立大学(North Carolina State University)的纳拉杨(J. Narayan)和保米克(A. Bhaumik)，通过高功率纳秒激光脉冲照射非晶碳



产生液态碳后迅速淬火而得到的。该物质根据温度的不同而分别具有半导体和金属性质。它还具有磁性，通过控制冷却速度还可形成微小的钻石。整个制造过程都在常温常压下，造价低廉，而且生产低成本的大块钻石是完全可能的，微小钻石则可用于智能显示器以及医药用途等。

产生液态碳后迅速淬火而得到的。

该物质根据温度的不同而分别具有半导体和金属性质。它还具有磁性，通过控制冷却速度还可形成微小的钻石。整个制造过程都在常温常压下，造价低廉，而且生产低成本的大块钻石是完全可能的，微小钻石则可用于智能显示器以及医药用途等。

(高凌云编译自2016年2月12日《欧洲核子中心快报》)