关于随机误差标准差的几点思考*

朱洪海
(盐城工学院 机械工程系,江苏 盐城　224003)

摘　要:从多种角度加深理解标准差的含义,有助于应用。用不同的代号表示贝塞尔公式和标准差的定义式不够妥当。

关键词:随机误差;标准差;统计均值;贝塞尔公式
中图分类号:TG801[TH124]　**文献标识码**:A　**文章编号**:1008-5092(2001)04-0020-02

在科学实验和工程实践中,普遍用到随机误差的标准差公式。较长时期以来,高校教材可能由于篇幅受到学时所限,往往对随机误差的标准差不作详细的论述,本文拟从多方面深入探讨标准差的含义,并就标准差的定义式和贝塞尔公式之间的关系作一讨论,希望与读者取得共识。

现约定本文所述随机误差均限正态分布,用表示随机误差,用表示随机误差的标准差。

1　关于标准差的含义

1.1　标准差的定义式

从数学上来说,随机变量 δ 符合标准正态分布时,其概率密度函数用下式表示:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\delta^2}{2\sigma^2}} \tag{1}$$

(1)式中的 σ 是正态分布曲线的特征参数,用下式表示:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \delta_i^{\,2}}{n}} \tag{2}$$

(2)式为标准差的定义式。具体到误差理论和数据处理中,σ 的含义要丰富得多。

1.2　标准差描述了随机误差的大小

1.2.1　标准差可以理解为随机误差绝对值的统计均值

对于同一个被测量,等精度测量结果的随机误差客观存在,但有大有小,有正有负,应该如何描述它的大小呢?任选其中一个 δ_i 显然不合适;取它们的算术平均值,也不合适,因为它们有正负相消性,使算术平均值趋近于0。所以只有取"平方和的平均值的平方根",不但避免了正负相抵消,而且"平均"其大小,又能获得和误差相同的量纲,因此,标准差可以理解为随机误差绝对值的统计均值。

1.2.2　n 个标准差的平方和等于 n 个真误差的平方和

将(2)式平方经整理后得

$$n\sigma^2 = \sum_{i=1}^{n} \delta_i^{\,2} \tag{3}$$

(3)式表示个相同的标准差的平方和,等于个不同的随机误差的平方和,这也可以理解成在数值幅度上是统计意义上的均值。

1.3　标准差体现了随机误差的分布特性

1.3.1　标准差体现随机误差的分散程度,反映测量准确度

正态分布曲线有两个特征参数,一个是变量的数学期望,它是 $n \to \infty$ 时变量的算术平均值;另一个就是 σ,表示变量对其算术平均值的分散程度,表示等精度测量结果的可靠程度,即表示测量准确度。对于同一个被测量,σ 小者,δ_i 的分布范围小,测量准确度高,测量结果的可靠性大;σ 大者,δ_i 的分布范围大,测量准确度低,测量结果的

* 收稿日期:2001-08-19
作者简介:朱洪海(1942-),男,江苏东台市人,盐城工学院副教授,主要从事机械设计制造精度及技术测量方向的教学与研究。

可靠性小。

1.3.2　体现随机误差统计意义上的平均状况

从概率积分的变量代换可知 $\delta = t\sigma$，t 为置信因子。当 t 取 ±1 时，$\delta = \pm\sigma$，这时的置信概率为 68.26%，也就是说有 2/3 以上数目的随机误差不超出 $\pm\sigma$ 的范围，占 δ_i 的大多数，另有将近 1/3 数目的随机误差超出了 $\pm\sigma$ 的范围。因而从统计角度看，$\pm\sigma$ 可以体现随机误差的平均状况，而且是偏于稳妥的平均状况，因为误差幅度小于 σ 的概率，等于误差幅度大于 σ 的概率的两倍以上。

1.3.3　$\pm 3\sigma$ 表示随机误差的分布范围

一般应用中，用 $\pm 3\sigma$ 表示的分布范围，3σ 则表示随机误差的最大幅值，置信概率达 99.73%。

综上所述，σ 不仅能描述随机误差统计意义上的数值大小，而且能描述随机误差的分布情况，所以称其为随机误差的“标准”差是很有道理的。

2　贝塞尔公式与标准差定义式的一致性

经以上分析可知标准差的定义式(2)是十分重要的。但是因为真值不知道，δ_i 无法一一算出，所以无法用(2)式求出 σ，需要应用贝塞尔公式即(4)式：

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \upsilon_i^2}{n-1}} \tag{4}$$

(4)式中 υ_i 为残差。这里需要强调的是：(2)、(4)两式等号左边必须用同一个代号 σ，两式中 ν 的含义和范围也是一致的。如果将(4)式等号左边换成和(2)式不同的其它代号，仍然叫做贝塞尔公式，就不够妥当了！理由如下：

2.1　关于贝塞尔公式的推证问题

(4)式所示的贝塞尔公式，不管是用“σ 是 δ_i 绝对值的统计均值”概念进行推导[1]，还是用样本的方差是总体方差的无偏估计来证明[2-3]，或是用其他方法推证，都要用到(2)式，(2)、(4)两式之间的关系是密切不可分割的，都要直接或隐含应用 $n\to\infty$ 作为条件，否则是推证不出贝塞尔公式的，这是符合随机误差正态分布大前提的。如果将(4)式等号的左边用与(2)式不同的其他代号表示，也将无法推证出贝塞尔公式。

2.2　贝塞尔公式与实验标准差[4]的表达式不可混同

不少教材和论著把(4)式改写成(5)式或类似(5)式：

$$S = \sqrt{\frac{\sum_{i=1}^{n} \upsilon_i^2}{n-1}} \tag{5}$$

(2)、(5)两式等号左边的代号已经不同，但在同一本书[4]（或论文）里仍然把(5)式称作贝塞尔公式，这是欠妥的，因为这时(5)式中的和(2)、(4)两式中的的范围已经不同。笔者的意见是：(5)式应当称作实验标准差表达式或样本标准差表达式，(5)式是贝塞尔公式的应用，而不是贝塞尔公式本身；为此，建议将等精度测量次数或样本大小作为的下角标，将(5)式改写为(6)式。和(5)式一样，(6)式是实验标准差表达式或称作样本标准差表达式：

$$\sigma_n = \sqrt{\frac{\sum_{i=1}^{n} \upsilon_i^2}{n-1}} \tag{6}$$

例如，某一被测量等精度测量 10 次的实验标准差记作 σ_{10}，等等。

把贝塞尔公式与实验标准差表达式混为一谈，其原因在于这些论著都是直接引用贝塞尔公式而未深究其来源，因而不发觉其中概念上的差别。在教学中，如果把两者混同了，就会感到因此而不能自圆其说。

2.3　测量次数对公式的影响

当时 $n\to\infty$，(2)、(4)两式一致，无需赘述。

实际工作中只能做到有限 n 次测量，这时候把(2)、(4)两式放在一起考虑也是一致的。理由如下：

(4)式的应用，通过(6)式即可算出实验标准差或样本标准差 σ_n，作为对理论标准差或总体标准差 σ 的估计值。因为是具体的有限 n 次测量，随机误差的分布已不是光滑对称的正态分布曲线，(1)、(2)两式都不是“这个”真实情况的反映。换句话说，(1)、(2)两式是“这个”有限 n 次测量的近似表达式。所以从有限次测量看，(2)、(4)两式也是一致的。所不同的只是应用(4)式可以得到 σ 的估计值，而(2)式却不能，这正是贝塞尔公式的高明之处，因而在科学实验和工程实践中得到广泛应用。但是，没有(2)式，就没有(4)式，所以(2)式有极重要的理论价值。

测量次数 n 影响对 σ 的估计准确度，对于同一个被测量，σ_{15} 的准确度显然比 σ_5 的准确度要高。

（下转第 28 页）

[7] 郭学锋,丁维平,颜其洁.一种新的制备纳米微粒的方法—快速均匀沉淀法[J].无机化学学报,2000,16(3):25~26.
[8] 岳林海,水淼,徐铸德.水解法制备的掺铁二氧化钛超微粉的热分析、晶体生长和结构相变[J].无机化学学报,2000,16(5):47~53.
[9] 张庆今,胡晓红,杨敏.液相沉淀法制备 TiO_2 超微粉末的影响因素分析[J].华南理工大学学报,1996,7:52~55.
[10] 全学军,李大成.制备高纯微细钛白粉过程中氯化铵的作用和影响[J].无机盐工业,2000,32(6):5~8.
[11] 毛日华,郭存济.液相反应制备纳米锐钛矿相二氧化钛[J].无机材料学报,2000,15(4):761~764.
[12] 李燕.水热法二氧化钛的制备[J].安徽建筑工业学院学报,1997,13:3~7.
[13] 高思勤.水热法合成纳米 TiO_2 及其在 Gratzel 电池中的应用[J].物理化学学报,2000,17(2):177~179.
[14] 吴风清.纳米 TiO_2 的制备、表征及光催化性能的研究[J].功能材料, 2001,32(1):69~72.
[15] 岳林海,刘清.纳米掺铁二氧化钛的 sol-gel 法制备与表征(1)[J].无机化学学报,2000,16(6):73~78
[16] 施利毅.微乳液反应法合成二氧化钛超细粒子[J].功能材料,1999,30(5):53~56.
[17] 崔正刚,殷福珊.微乳化技术及应用[M].北京:中国轻工业出版社,1999.
[18] 张志琨,崔作林.纳米技术与纳米材料[M].北京:国防工业出版社,2000.
[19] 张立德,牟季美.纳米材料和纳米结构[M]. 北京:科学出版社,2001.

Progress in Study on Nanometer Titanium Oxide Prepared By liquid-phase Reaction

WU Ying-chun, XU Zhang-fa
(Dept. of Environmental Science, East China of Normal University　200062, PRC)

Abstract: Nanometer titanium dioxide can be made by gas - phase method and liquid - phase method. The later is focused on for its benefits. Current research progress in liquid - phase reaction preparation processes and characterization methods of nanometer are reviewed in the paper. Suggestions on research orientation are also raised.

Keywords: liquid-phase method; nanometer titanium dioxide; study; preparation; characterization

(上接第 21 页)

3 结论

(1)标准差不仅是随机误差的分布特征参数,而且标准差可以理解为随机误差绝对值的统计均值。标准差能够比较全面地描述随机误差。

(2)标准差的定义式和贝塞尔公式是一致的,应当用同一个代号表示它们。贝塞尔公式是标准差从理论过渡到实际应用的桥梁,应当注意区分贝塞尔公式与实验标准差表达式或样本标准差表达式之间的差别。

参考文献:

[1] 朱洪海.对贝塞尔公式证法的探讨[J].计量与测试技术,2001,(6):8~9.
[2] 刘次华,万建平.概率论与数理统计[M].北京:高等教育出版社,1999.
[3] 施雨,李耀武.概率论与数理统计应用[M].西安:西安交通大学出版社,1998.
[4] 国家质量技术监督局计量司.测量不确定度评定与表示指南[M].北京:中国计量出版社,2000.

Thinking of Standard Error of Random Error

ZHU Hong - hai
(Department of Mechanical Engineering of Yancheng Institute of Technology, Jiangsu Yancheng　224003, PRC)

Abstract: Further understanding meanings of standard error in various ways is useful to application. And it is unsuitable to use different letters to indicate the Bessel formula and the definition formula of standard error.

Keywords: Random Error; Standard Error; Statistical Mean; Bessel Formula