

第 1 篇

实验数据分析与经验证据评估

清华大学理学院前院长周光召教授曾说：“为什么 20 世纪最重要的物理学发现又恰恰在德国的土地上发生？首先德国人非常重视实验和实验数据的分析。从普朗克开始分析黑体辐射到后来原子物理中玻尔提出他的原子论，最关键的就是对光谱的分析。当时德国对光谱的分析可能是最多的，包括海森伯就是从光谱分析而提出矩阵力学的。他们的理论是和实验密切地结合在一起的，这是德国物理最大的一个特点。第二个特点就是德国有很强的数学传统。当时德国尽管其他学科不怎么发达，数学已经超过英国了，在 20 世纪，德国就有黎曼、高斯、希尔伯特，本世纪初就成为世界的第一位，哥廷根一直是世界的数学中心。第三个特点是德国有很强的哲学传统。这三个条件：理论紧密地和实验结合在一起，非常强的数学传统和打破哲学上的机械论，对于德国能在这种环境下产生 20 世纪最伟大的科学发现起了决定性的作用。”^[1.01]

科学是系统的知识体系，科学是探索自然的过程，科学也被看作一种社会建构^[1.02]。虽然不同学派的定义不同，但普遍认为科学的适当目的是追求对理论的说明，“达到越来越深刻、越来越统一和越来越有预见力的理论”^[1.03]。科学追求包括数学严密性在内的逻辑自洽性，寻求主要源于科学实验的可重复或可比对的经验证据的支持^[1.04]。对实验提供的只是检验理论的观察、陈述的观点有两类批评，其中“The most serious philosophical objection is that logical positivism and post-positivist philosophies of science have undermined experiment’s traditional, privileged position as a source of empirical knowledge”^[1.05]¹²⁰，实验证据相对于理论的独立性也被质疑^[1.06]³⁸，但实验毕竟是感性经验的基本通道，是经验知识的重要源泉（引自曾国屏教授对作者的来信），不能轻视实验在科学探索中的重要性。实验不仅能对理论命题提供一定程度的可重复经验证据支持、或对否认提供有条理的质疑参考。而且，“As a knowledge-producing activity, experiment engages the inchoate, the practical, and the particular. ... An essential role of experimentation is to provide new

information about how to investigate the world; in other words, instrumental knowledge about the world under investigation.”“在当代科学中,运用探索性、非论证的实验依旧是重要的。”^[1.05]¹¹⁸实验常常具有“它自己的生命”^[1.05]¹²⁰。

《Science》杂志主编 Alberts B. 2011 年 1 月撰文批评以往的大学科学教育说:“……作为一名主修科学的本科生,……头两年里,我每周花三个下午在这种实验课上。同学中后来成为科学家的大多数人,都认为这些实验课是冗长无味的‘炒菜课’,……所学的就是紧跟指引、照章操作。诚然,我们的确接触了不少仪器设备,比如测量重量和液体的器件;我们还学习了如何做实验笔记。但是我们既没有掌握对科学本质的真正理解,也没有获得关于产生和评估科学证据及其解释的任何经验,而这两点才是现代意义下‘科学教育’最核心的要素。许多学校至今仍然一成不变地保留着这种缺憾。”^[1.07]

剑桥出版社 2003 年出版的 G. L. Squires 编写的第 4 版《Practical Physics》中,数据统计处理与记录计算两部分占 195 页中的 120 页^[1.08]。

密歇根大学的 Byron P. Roe 1992 版《Probability and Statistics in Experimental Physics》一书约 200 页,前言中强调该书“is meant to be a practical introduction into the use of probability and statistics in experimental physics for advanced undergraduate students and for graduate students.”^[1.09]

S. L. Meyer 的《科技工作中的数据分析》一书 1983 版中译本有 78 万字,分 5 部分,第 1、2 部分曾被作为美国西北大学初级物理实验补充读物,第 3、4 部分纳入了物理和化学系的教学大纲^[1.10]。

美国物理教育家 H. F. Meiners 等在《Laboratory Physics》第 2 版的前言中说:“The objectives of this laboratory are: 1. To introduce the student to the significance of experimental approach through actual experimentation. 2. To apply the theory of the textbook and the recitation class to real-life problems to develop a better understanding of the fundamentals of classical and modern physics. 3. To introduce the student to the methods of data analysis used throughout science and engineering. 4. To develop an error conscience”,使工程师和科学家应该知道她的测量的相对价值,而不管其测量类型如何。其第 1 版中译本数据处理导论部分约 8 万字^[1.11]。

赵忠尧在《实验物理中的概率和统计》一书的序中说:“科学观测或实验……实验数据的获取和实验数据的分析处理是缺一不可的两大环节”^[1.12]。这也就是 Bruce Alberts 所说“关于产生和评估科学证据及其解释”的两个环节。

本篇介绍的就是“评估科学证据”所涉及的重要基础概念和经验方法。可以说:分析处理实验数据,是寻求可重复或可比对经验证据过程的重要步骤,是追求逻辑自治性与数学严密性的过程,也是科学理性精神的实践体现环节。

数据分析与不确定度评定基础

1.1 物理量与测量

科学实验离不开对现象的观察,也离不开对表征状态或过程的物理量的测量。

国际计量局在《A concise summary of the International System of Units (8th edition, 2006)》的页首写道:“Metrology is the science of measurements,made at a known level of uncertainty,in any field of human activity.”^[1.13]这说明了测量的普遍性和重要性,也说明了在测量结果中评定不确定度的必要性。物理实验中,定性和半定量观测固然重要,但基于对物理量科学测量的定量研究通常更为重要。

国际计量学通用名词术语《VIM 2007》中定义:测量是用实验方法获得量的量值的过程。可测量的量(measurable quantity)是“现象、物体或物质的可以定性区别和定量确定的属性”^{[1.14]16}。

在物理实验中,不仅要明确测量对象,选择恰当的测量方法,正确完成测量的各个步骤,还要学习误差理论和实验数据处理的基本概念,学会能够对多数测量表示出完整的测量结果,包括表示出确定置信水平的不确定度。

目前国内物理实验教材中在不确定度分量评定与合成方法上主要靠拢两类模式:(1)参考 ISO/IEC《测量不确定度评定指南》^[1.15]和相应国家计量技术规范^[1.16]的模式,其关键特征是先评定标准差形式的两类不确定度分量,再求方和根得出合成标准不确定度,简称为《指南》模式、GUM 模式或 u_c 合成模式;(2)参考中国计量科学研究院(NIM)2002 年建议信的模式^{[1.17]94-96},借鉴经典的误差理论的研究成果^[1.18,1.19],其关键特征是先评定出高置信概率的两类不确定度分量,再求方和根得出合成扩展不确定度,简称 NIM 建议信模式、极限值合成模式、“经典”^[1.14]模式或 U 合成模式。清华大学教学中我们最新提出的不确定度合成的技术综合方法(THM)及简化程序见第 4 章,计量学与仪器仪表科学中近年开始推行的不确定度合成的蒙特卡罗方法(MCM)见第 3 章。

u_c 合成模式 在部分物理实验教材基础上 作了补充修改 ^[1.20,1.21]	U 合成模式 在清华大学物理实验教程基础上 作了增补 ^[1.22,1.23]
[例 1] 用三个 0.1 级电阻箱组成自组电桥测某个电阻,测量结果写成下式 $Y = y \pm u_c = (4030.0 \pm 4.5) \Omega$ $(\nu_{\text{eff}} \approx 78) \quad (1.1)$ 式(1.1)中 $u_c = 4.5 \Omega$ 是合成标准不确定度(combined standard uncertainty). 式(1.1)表示: y 与被测量的真值 Y_t 之差 $y - Y_t$ 大部分位于区间 $(-u_c, u_c)$ 之内,位于该区间内的置信概率约为 $2/3$ (不等于 $0.683^{[1.17]65}$). 提高要求: 同时写出有效自由度 ν_{eff}, 它被用来计算扩展不确定度, ν_{eff} 的意义见 1.4 节。	[例 1] 用三个 0.1 级电阻箱组成自组电桥测某电阻,测量结果写成下式 $Y = y \pm U = (4030 \pm 9) \Omega$ $(p \approx 0.95) \quad (1.1')$ 式(1.1')中的 $U = 9 \Omega$ 是置信概率约等于(或大于)0.95 的扩展不确定度(expanded uncertainty)。式(1.1')表示: y 与被测量的真值 Y_t 之差 $y - Y_t$ 位于区间 $(-U, U)$ 之内的置信概率(可能性)约 95%。教学中约 0.95 的置信概率说明可不写出。各国工业技术规范标准,特别是 ISO9000 系列的“测量管理体系”标准中,大多数要求结果表示具有约等于或大于 0.95 的置信概率。

完整的测量结果表示中,必须包括测量所得的被测量(measurand)的数值($y/\Omega = 4030$)和测量单位(Ω),一般应给出不确定度。必要时还需写出对测量结果有作用的影响量的值。如测电流计内阻时写出室温 $t = (20.0 \pm 1.0)^\circ\text{C}$, 因为内阻值随温度升高而增加,温度是重要的影响量。测量对象、测量单位、测量方法和测量不确定度曾被称为测量的四个要素。

注 1.1 置信概率与概率

数据统计分析中,与置信区间有关的置信概率也称置信水平,它与概率论中的一般概率概念有所不同,通俗表述的“约定真值 Y_t 落在区间 $(\bar{x} - U, \bar{x} + U)$ 之内的概率”是不够严密的^[1.24], 因为 Y_t 有确定的值,而置信限 $\bar{x} \pm U$ 是随机变量。早期文献强调表述置信概率时的“约”字不宜省略^{[1.25]15}。考虑到一些 B 类不确定度分量来源给出值往往包含“足够可靠”的含义,按钱钟泰的建议增加了“或大于”的限定词^{[1.22]4}。

1.2 误差的定义、分类及简要处理方法

1.2.1 测量误差的定义

通常说误差是测量结果 y 和被测量的真值或约定真值 Y_t 之差 dy 。

$$dy = y - Y_t \quad (1.2)$$

注 1.2 误差的新定义

文献[1.14]第 22 页中,“error”更普遍的定义是“measured quantity value minus a reference quantity value”。该文本条目的注释中指明了参考值常常是约定的真值。参考值一词的含义更为广泛,能为不同领域、不同学派广泛接受,因为某些科学哲学流派排斥“真

值”的概念。这里误差用微分符号表示,为强调说明误差是小量、且可正可负。

真值是理想的概念,只有定义严密时通过完善的测量才可能获得或接近之,它一般无从得知。因此一般不能计算误差,只在少数情况下用准确度高的实际值作约定真值时才能计算误差。

[例 2] 用准确度等级为 1.5、量程为 5mA 的电流表测精密恒流源 2.0000mA (约定真值)的输出电流,估读间隔为 0.02mA,三次读数分别为 2.02mA、2.00mA 和 1.98mA,示值误差分别为 0.02mA、0.00mA 和 -0.02mA。

误差的普遍性 由于测量仪器不准确、原理或方法不完善、环境条件不稳定、人员操作不熟练等原因,任何测量结果都可能具有误差。虽然因一般不知道真值而不能计算误差,但是能分析误差产生的主要因素,能减小或基本消除有的误差分量对测量的影响。对测量结果中未能消除的误差影响,要估计出它们的极限值或表征其分布特征的参量,如标准偏差。误差的普遍性要求我们:必须重视对测量误差的分析,重视不确定度评定,尽可能完整地表示测量结果。

1.2.2 误差的分类及简要处理方法

误差主要分为两类:随机误差和系统误差。它们的性质不同,应分别处理。

注 1.3 误差分类的补充:粗大误差

还有一类误差,由于外界干扰、操作读数失误等原因而明显超出规定条件下的预期值,以前俗称粗大误差或粗差。包含粗大误差的测得值或粗大误差称为统计离群值(statistical outlier, outlier),曾被称为高度异常值^{[1.22]108}。测量要避免出现统计离群值。已被谨慎确定为统计离群值的个别数据可以剔除,详见第 2 章。

1.2.2.1 随机误差(random error)

(1) 随机误差的定义

随机误差是重复测量中以不可预知方式变化的测量误差分量。电表轴承摩擦力矩的变动、螺旋测微计的测头压紧力在一定范围内变化、操作读数时在一定范围内随机变动的视差影响、数字仪表末位取整数时的随机舍入过程等,都会产生一定的随机误差分量。

[例 3] 用惠斯通电桥测微安表内阻时,需降低电源电压以避免过载。若还用原配的检流计,电桥将变得不灵敏。检流计的灵敏阈一般取偏转 0.2 分格对应的量值,它近似为可觉察的最小电流改变。电桥的比率臂值为 1 时,如平衡后要使检流计平均偏离 1 分格,被测电阻需改变 δ_R ,近似可得为偏离 0.2 分格,电阻需要改变 $0.2\delta_R$,那么由灵敏阈决定的测量误差 $e(R)$ 是随机误差分量,它在区间 $\pm 0.2\delta_R$ 内时大时小,时正时负。

注 1.4 误差分量或不确定度分量的十项来源

定义不严密、复现法不理想、环境条件不稳定、取样代表性不够等,会使被测量的体现值波动,被测对象有“原分布”,这些影响也反映在测得值波动中,细析不确定度来源时需考虑。在对某教材的修改建议中,笔者对文献[1.15]中不确定度来源的十个方面作了改写与重排,也可看作是误差分量的十项来源:①被测量的定义不完善;②相同条件下被测量在重复观测中的变化;③复现被测量的方法不理想;④取样的代表性不够;⑤测量方法和测量程序的近似和假设;⑥测量仪器的(测量不确定度、有效分辨率、鉴别力阈等)计量性能局限;⑦测量标准或标准物质的不确定度;⑧引用数据或其他参量的不确定度;⑨对主要环境条

件等影响量的认识不当或控制、测量不完善；⑩仪器读数有人为偏移，测量操作有人为偏差，计算模型、程序方法或数据修约有不恰当的简化或人为疏误。10条中，第①、②条重点关注被测量；第③、④、⑤条关注获取（反映被测量的）测量值的具体方法和程序；第⑥～⑨条关注测量设备及测量条件，在本世纪 ISO 9000 系列的新国际标准[1.27]中，广义的测量设备包括测量标准、标准物质、测量规范文件及引用数据；第⑩条突出人为因素的影响。文献[1.28]中对几何量测量的十项来源作了更详细的分解讨论。

随机误差分量是测量误差的一部分，其大小和符号虽然不知，但在相同条件下对同一稳定被测量的多次重复测量中，它们的分布常常满足一定的统计规律。随机误差分布绝大多数是“有界性”的，大多数有抵偿性，相当多的有单峰性，即绝对值小的误差出现的概率较大。

(2) 算术平均值 (arithmetic mean or average)

大多数随机误差有抵偿性，即测量次数足够多的时候，正、负误差之和的绝对值近似相等。因此用多次测得值的算术平均值作被测量的估值，能减小随机误差的影响。一般应使测量次数 $n \geq 6$ 。如果对同一量重复测了 n 次，测得值为 y_i ，平均值为

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (1.3)$$

(3) 实验标准[偏]差 (experimental standard deviation)

随机误差引起测得值 y_i 的分散性用实验标准偏差 s 表征，由贝塞尔法 (Bessel method) 算出

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (1.4)$$

s 反映了随机误差的分布特征。 s 大表示测得值分散，随机误差的分布范围宽，精密度 (precision) 低； s 小表示测得值密集，随机误差的分布范围窄，精密度高。

测量准确度 (accuracy) 的定义为 closeness of agreement between a measured quantity value and a true quantity value of the measurand^{[1.14]21}。准确度反映随机误差和系统误差的综合影响程度，它是一个定性的概念。测量准确度的定量表征用测量不确定度或测量误差限值。不确定度小，准确度高。

注 1.5 慎用“精度”一词

准确度曾叫“精度”，目前仅少数几何量专业的教材等坚持延续用“精度”一词，不具规范性。现在规范中精度只是精密度的简称。在口语和书面文字中要尽量避免使用“精度”一词。但有一类特例：在加权平均或加权回归时，一般要求测得值或因变量值等精密度，俗称等精度，这里的“等精度”原来仅指“以标准偏差为表征值的精密度大致相等”，但实用中常常推广到“与扩展不确定度（或误差限）为表征值的准确度大致相等”。

求标准不确定度 u_c 时用到平均值的实验标准偏差 $s_{\bar{y}}$ (experimental standard deviation of the mean)

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (1.5)$$

注 1.6 平均值的实验标准偏差不能用于方法、仪器或传感器的特性表征

少数论文中用平均值的实验标准偏差 $s_{\bar{y}}$ 来表征其试验装置或方法的特性，并据此宣称达到“国际最高水平”，这是违背科学的做法。对于测量传感器、方法或仪器，在同一被测量

的多次重复测量中,当测量次数 n 趋于 ∞ 时。表征其精密度(测得值分散性)的实验标准偏差 s 趋于总体标准差 σ 。在现代检测技术中增加次数 n 很容易, $s_{\bar{y}} \approx \sigma/\sqrt{n}$, 显然有 $\lim_{n \rightarrow \infty} s_{\bar{y}} \approx 0$ 。有的学人在明知这一趋势时还宣称其精密度指标“国际领先”,是学风不端的表现。对此问题,国内外标准文献[1.29~1.32]均已发布规范性论述。

(4) 残差与最小二乘法

残差(residual, residual error)是测量列中某一测得值 y_i 与该测量列的算术平均值之差。更一般的定义为 y_i 与其(最佳)估计值 $\hat{y}_i$ 之差,记作

$$v_i = y_i - \hat{y}_i \quad (1.6)$$

最小二乘法(least squares method, LSM)。数据处理要充分利用测量所获得的有效信息,以减小误差对结果的影响。LSM 是一种根据实验数据求未知量“最佳”估值的方法,最小二乘原理表述为:使(等精密度的因变量) y 的残差平方和(residual sum of square, RSS)或标准差的平方(方差)为极小值^[1.33]。残差平方和最小记作^[1.32]

$$RSS = \sum v_i^2 = \min \quad (1.7)$$

1.2.2.2 系统误差

系统误差(systematic error)是重复测量中保持恒定或以可预知方式变化的测量误差分量,简称系差。

系统误差举例:指针式电表的零位误差,电表分度或磁场分布不匀,伏安法测电阻时电流表内接或外接引起的表内阻影响的分量,单摆运动方程小角度解近似,螺纹副的螺距有误差,测量时温度等影响量对额定值的偏离,空气浮力对天平质量称量的影响,空气折射率近似地取 1 等,都会产生一定的系差分量。

系统误差包括已定系差和未定系差。

(1) 已定系差

指符号和绝对值已经知道的误差分量。实验中应尽量消除已定系差,或对测量结果进行修正,得到已修正结果(corrected result)。修正式为

$$\text{已修正测量结果} = \text{测得值(或其平均值)} - \text{已定系差} \quad (1.8)$$

修正值(correction)等于负已定系差。已修正结果等于测得值加修正值。

(2) 未定系差

指符号或绝对值未被确定而未知的系差分量。一般只能估计其限值或分布特征值。未定系差分量大多和下文的 B 类不确定度分量来源有大致对应关系。对于不同测量条件、不同被测量值或不同时段等,未定系差在一定意义上可以说具有随机性。例如在 $(20.0 \pm 2.0)^\circ\text{C}$ 的空调室内,某一时刻室温对 20.0°C 的偏离误差是定值系差,但不同时刻的偏离误差在 $\pm 2.0^\circ\text{C}$ 内变动,变动范围已知,但分布规律未知,具有随机性。变频空调的温度偏离误差分布常有单峰性,到达温控极值点就启停的开关式空调偏离误差分布一般无峰。

误差的随机性,包括随机误差的随机变量特性和未定系差的某种“随机性”,是不确定度分量方和根合成法的基础,因而也是整个误差分析与不确定度评定体系的重要客观依据。

注 1.7 误差的随机性是数据处理的重要基础

随机误差分量的严格“随机性”,是经典误差理论的基础。未定系差的近似“随机性”,是

现代误差理论中误差表征值标准差合成(综合)、相近水平的误差限方和根合成的基础。在不确定度评定体系建立初期的1980年,“B类标准不确定度”的约定定义、汇总各国意见后约定采用的不确定度分量合成的“方和根法”^[1.15],就是以未定系差的一定随机性为重要依据。因此,实验数据分析与不确定度评定,并不具有严密的公理化体系。不同学派的学者都公认“误差方差的运算定律”(等效于标准差分量合成的方和根法),是一系列规范赖以建立的“核心理论基础”^{[1.18]1}。在不确定度体系创建初期,BIPM征询了32个国家计量实验室及5个国际组织的意见,收到21个国家实验室的回复,方和根合成法就是综合大多数回复意见后的统一约定^{[1.33]214,[1.34]}。

系差分析的重要性 大量一般测量的实践表明,除了少数比较测量之外,系统误差分量对测量结果的影响常常显著地大于随机误差分量的影响。因此大学物理实验要重视对系差的分析,尽量减小它对测量结果的影响:①对已定系差进行修正;②合理评定系差分量大致对应的B类不确定度分量;③通过方案选择、参数设计、计量器具校准、环境条件控制、计算方法改进等环节来减小系差的影响。物理实验中,经常用多个散布测量点作直线拟合求斜率,将一些未定系差随机化,是减小系差影响的方法之一。

1.3 直接测量结果的不确定度评定

不必测量与被测量有函数关系的其他量,就能直接得到被测量值的测量方法叫直接测量法。例如用等臂天平测质量、用电流表测电流等都是直接测量。

用式(1.1)形式表示的直接测量结果中,不计已定系差时,被测量值 y 可取多次测量的平均值 $\bar{y}$;若只测一次, y 就取单次测得值 y_1 。如有已定系差,还须按式(1.8)将测得值或其平均值减去已定系差,得到 y 的值。

1.3.1 测量不确定度的概念及其与误差、误差限的关系

“Uncertainty of measurement: non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used.”^[1.35]

不确定度是与测量结果相关联的参量,是表征被测量的真值(或指与定义、测量任务相关联的被测量值,或指参考值)所处的量值散布范围的评定。

u_c 合成模式	U 合成模式
[例4] 例1中电桥法测量电阻的结果为$\hat{R}_x \pm u_c = (4030.0 \pm 4.5)\Omega$,有效自由度$\nu_{\text{eff}} \approx 78$,说明电阻量值4030.0的误差在区间$\pm 4.5$之内的置信概率约2/3。如将其其他方法测得的$u_c = 0.05$的值4033.00$\Omega$作约定真值,电桥法实验值4030.0的误差为$-3.0\Omega$。	[例4] 氢光谱实验中,测得H_α线15$^\circ\text{C}$时的波长为$\lambda_{H_\alpha} = (656.24 \pm 0.06)\text{nm}$,测量不确定度为0.06,说明波长真值一般在656.18~656.30nm之间,量值656.24的误差一般在$-0.06 \sim 0.06$之间。将文献中的高准确度测得值656.2816作约定真值^[1.36],实验值656.24的测量误差为-0.04nm。

不确定度表示由于测量误差的存在而对被测量值不能确定的程度。不确定度反映了可能存在的误差分布范围,即随机误差分量和未定系差分量的联合分布范围。

用合成标准不确定度 u_c 表示时, u_c 正比于一定置信概率的误差限(本书中指绝对值)。可近似说:误差在区间 $(-u_c, u_c)$ 内的概率约为 $2/3$ ^{[1.17]65} 。	用扩展不确定度 U 表示时, U 为误差分布基本宽度的 $1/2$ 。误差一般在 $\pm U$ 之间,在 $\pm U$ 外的置信概率不大于 5% 。
---	---

不确定度体系吸收、借鉴了现代误差理论的研究和实践成果。近似地说,不确定度大致对应于一定置信概率的误差限的绝对值。不确定度总是不为零的正值,而误差可能为正,可能为负,也可能十分接近于零^[1.15]。单位与修约间隔确定的检定报告中,具体的检定值的误差可出现零,例如 0.5 级 50mA 电表的检定证书中,某些整刻度的示值误差可写成 0.00mA。不确定度总是可以具体评定的,而误差一般因真值未知而不能计算。

实验数据处理时,通常先作误差分析(现在也称为不确定度分量来源分析),必要时谨慎地剔除统计离群值,接着再修正已定误差,然后评定不确定度。

注 1.8 对不确定度与误差、误差限之间关系的补充讨论

(1) 扩展不确定度与误差、误差限的区别表

讨论主体	计算基点	基本含义	一般特征
误差 Δy	约定真值 Y_{tr} (参考值)	测得值—真值	可正可负 一般不能计算
误差限 $E_{\pm lim}$	约定真值 Y_{tr} (参考值)	一定置信概率 p 时(最佳估值 $Y_{op\ ev}$ 减去约定真值 Y_{tr} 所得的)误差分布区间的极限值	有正有负 能计算
	误差限相对于约定真值 Y_{tr} $Y_{op\ ev} - Y_{tr} = - E_{-lim}$ $Y_{op\ ev} - Y_{tr} = E_{+lim}$ $P(- E_{-lim} \leq Y_{op\ ev} - Y_{tr} \leq E_{+lim}) \approx 95\%$		
不确定度 U	修正后的平均值等最佳估值 $Y_{op\ ev}$	一定置信水平 $1-\alpha$ 时被测量值 Y 相对于最佳估值 $Y_{op\ ev}$ 的散布区间的半宽度	不为零的正值 能计算
	相对于最佳估值 Y_{op}	$P(Y - Y_{op\ ev} \leq U) \approx 1 - \alpha \approx 95\%$ 或大于	

(2) 误差(限)相对于约定真值 Y_{tr} 。评定误差时往往不包括、不考虑被测量定义、复现不完善或体现值波动的影响。研究被测量值、评定不确定度时,一般要考虑被测量定义、复现不完善或体现值波动的原分布影响^{[1.12]145-147}。

(3) 误差分量和不确定度分量在分析处理时往往不对应,有交错或互含:随机误差影响的分量,常常用统计学方法处理,作为不确定度的 A 类分量;但不少 B 类分量中也可能包含他人已经用统计学方法处理的部分,例如一些仪器文件资料给出的单一的 MPE 中,就已经包含了部分这样的成分。再如线纹尺与被测线段的夹角 θ_i 在 0° 附近呈随机误差分布时,对线段长度测量的影响称为余弦误差(投影误差)^[1.37],它的相对值为 $1 - \cos\theta_i \leq 1$,余弦误差中只有一部分具有抵偿性因而能用统计学方法计算 A 类不确定度分量,对不具有抵偿性的部分要作 B 类不确定度分量的估计。

(4) “不确定度”取代“误差”的原因有:①现代误差理论已有很大的发展。但“误差”与表征误差分布特征的加定语的概念“ $\times\times$ 误差”之间常被混淆。②普遍认为:对某组测量结果,即使所有已知的或怀疑存在的误差分量已经被估计并作了适当的修正,仍然会存在某种不确定度^[1.15]。③非统计学方法估计的分量愈显重要。分量合成方法长期不统一,给国际比对与技术商贸交流带来困难,亟待概念与评定方法一致、可比对的新规范、新方法。④误差体系的分类与合成法有不完善或瑕疵,如 25mm 一级千分尺示值误差限为 $4\mu\text{m}$,常作为系统误差来考虑。但在 $4\mu\text{m}$ 中估读误差、瞄准误差、视差影响、压陷量变动等小的分量主要属于随机分量,一级螺纹副的制造公差 $\pm 3\mu\text{m}$ 对应的误差属系统性误差分量^[1.38]。⑤20 世纪科学哲学众多学派的影响,如在对“真值”、“客观”、“实在”等概念上不同学派有着不同的理解,因而在不确定度体系与现代误差理论的关系上存在分歧。⑥少数学者力图建立“准”公理化的不确定度体系以完全取代现代误差理论。目前已建立并在不断完善的不确定度体系表面看来系列概念似乎是“严密的”,但在方和根合成法及其他评定方法中,存在着几条非公理化的规定或约定建议(详见第 4 章)。从本质上说,不确定度体系和现代误差理论都不是公理化体系,它们虽然都以统计学为理论指导和重要基础,但也都以实验测量的客观实在为重要出发点。这里所说的客观实在包括测量的实际情况、实际规律、实际需求、实际存在的规范标准或普遍性约定。

1.3.2 不确定度的简化评定方法

依据国家计量技术规范和国内外一些标准、规范的精神,参考《指南》或 NIM 建议信,教学中采用比较简化的、比较适用于一般测量的不确定度评定方法。要点如下。

u_c 合成模式	U 合成模式
1) 结果表示采用合成标准不确定度 u_c 合成标准不确定度 u_c 用于测量结果的报告。式(1.1) $Y = y \pm u_c$ 表示被测量量值位于 $y \pm u_c$ 内的置信概率在有效自由度 ν_{eff} 大时约为 2/3。结果报告中 u_c 一般需附 ν_{eff} , 详见下文。	1) 结果表示采用扩展不确定度 U 扩展不确定度 U 用于测量结果的报告,也称报告不确定度。式(1.1') $Y = y \pm U$ 表示被测量量值位于区间 $y \pm U$ 内的置信概率约等于或大于 0.95。扩展不确定度有时简称不确定度。