

中华人民共和国国家标准

GB/T 27418—2017

测量不确定度评定和表示

Guide to the evaluation and expression of uncertainty in measurement

(ISO/IEC Guide 98-3:2008, Uncertainty of measurement—
Part 3: Guide to the expression of uncertainty in measurement, MOD)

2017-12-29 发布

2018-07-01 实施

中华人民共和国国家质量监督检验检疫总局 发布
中国国家标准化管理委员会

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准使用重新起草法修改采用 ISO/IEC 指南 98-3:2008《测量不确定度 第 3 部分:测量不确定度表示指南》。

本标准与 ISO/IEC 指南 98-3:2008 相比没有技术性差异,但在结构上有部分调整,附录 H 中列出了本标准与 ISO/IEC 指南 98-3:2008 的章条编号对照一览表。

本标准还做了下列编辑性修改:

- 标准名称改为“测量不确定度评定和表示”;
- 删除了国际标准中部分关于 2008 版文件变化的脚注;
- 删除了国际标准中引用的法文内容;
- 将国际标准第 4 章的第 1 段“关于评定不确定度分量的实用指南可参见附录 F”调整为该章的最后 1 段;
- 对国际标准中部分示例中的数值做了修约;
- 本标准用“包含区间”和“包含概率”替代国际标准中的“置信区间”和“置信水平”;
- 修改了国际标准“B.2.1.1 被测量”的内容,增加了术语“被测量”在 VIM 第 3 版和 VIM 第 2 版中的差异的说明;
- 附录 B 增加了 B.1.4.9,说明术语“测量结果”在 VIM 第 3 版和 VIM 第 2 版中的差异;
- 附录 G 的表格增加了栏目名称;
- 删除了参考文献中的脚注,原脚注中的内容调整到对应的条目下以注的方式给出;
- 参考文献中增加了 2 个国内相关文献。

本标准由全国认证认可标准化技术委员会(SAC/TC 261)提出并归口。

本标准起草单位:中国合格评定国家认可中心、中国计量科学研究院、天津出入境检验检疫局、北京理工大学、中国船舶重工集团公司第七一五研究所、北京长城计量测试技术研究所、中国测试技术研究院。

本标准主要起草人:张明霞、林志国、吉黎明、施昌彦、王阳、安平、王春艳、贾小川、陈凌峰、陆渭林、崔岩梅、任冬梅、刘浩峰。

引 言

0.1 当报告物理量的测量结果时,应对测量结果的质量给出定量的说明,以便使用者能评价其可靠程度。如果没有这样的说明,测量结果之间不能进行比较,测量结果也不能与标准或规范中给出的参考值进行比较。所以需要有一个便于实现、容易理解和公认的方法来表征测量结果的质量,本标准给出的就是一个这样的方法,即评定和表示其测量不确定度。

0.2 虽然误差和误差分析早就成为测量科学或计量学的一部分,但作为定量特征的不确定度概念还是一个比较新的概念。现在大家都认识到,当对已知的或可疑的误差分量都作了评定,并进行了适当的修正后,这样的测量结果仍然还存在着不确定度。

0.3 正如国际单位制(SI)已在全世界的所有科学和技术测量中普遍使用一样,全世界对测量不确定度的评定和表示方法取得共识,给出易于理解和有恰当释义的规则,将会对科学、工程技术、商贸和工业中大量的测量结果具有极为重要的意义。在市场全球化的时代,推广和应用统一的不确定度的评定和表示方法势在必行,以便使不同国家进行的测量可以容易地相互比较。

0.4 评定和表示测量结果不确定度的理想方法应该具有:

——普遍适用性:方法应该适用于所有类型的测量和测量中用到的各种输入量。

用于表示不确定度的实际的量应该是:

——内部协调的:它应该直接由对它有贡献的分量导出,且与这些分量如何分组以及这些分量如何分解成子分量均无关。

——可传递的:如果第二个测量中使用了第一个测量的结果,则应该可以用第一个结果的不确定度作为评定第二个测量结果的不确定度的一个分量。

此外,在许多工业、商业以及健康和安全管理领域中,常常有必要提供测量结果的一个区间,可期望该区间包含了被测量之值合理分布的大部分。评定和表示测量不确定度的理想方法应能方便地给出这样一个区间,特别是在符合实际需要的包含概率或置信水平下的区间。

0.5 本标准提出的方法是以建议书 INC-1(1980)^[2]给出的纲要为基础的,该建议书是根据 CIPM 的要求,由 BIPM 组建的不确定度声明工作组起草的。该方法的合理性在附录 C 中进行了讨论,它满足了上述全部要求。现在使用的大多数其他方法则并非如此。建议书 INC-1(1980)是由 CIPM 批准,并在其建议书 1(CI-1981)^[3]和建议书 1(CI-1986)^[4]中再次确认;附录 A 中的 A.2 和 A.3 分别是 CIPM 建议书的 CI-1981 和 CI-1986。由于建议书 INC-1(1980)是这些文件的基础,所以在本引言的 0.7 中全文给出了它的内容。INC-1(1980)建议书权威的法文本参见 ISO/IEC 指南 98-3:2008 的 A.1。

0.6 本标准的第 8 章概括介绍了评定和表示测量不确定度的程序,并在附录 F 中详细介绍了一些范例。此外,本标准还有其他一些附录,如附录 B 是“真”值、误差和不确定度及其他概念;附录 D 是评定不确定度分量的具体建议;附录 E 是自由度和置信水平(包含概率);附录 G 为本标准用的基本数学符号;以及参考文献目录。

0.7 建议书 INC-1(1980)《实验不确定度的表示》

1) 测量结果的不确定度通常包含若干个分量,根据其数值的评估方法不同分为两类:

A 类:用统计方法评定的分量;

B 类:用其他方法评定的分量。

A 类和 B 类不确定度与以前所用的“随机效应引入的不确定度”和“系统效应引入的不确定度”之间并不总是存在简单的对应关系。“系统不确定度”这个术语容易引起误解,应该避免使用。

任何有关不确定度的详细报告应该有一个完整的分量明细表,对每个分量应该说明其数值的获得方法。

- 2) A类分量用估计方差 s_i^2 (或估计标准差 s_i) 及自由度 ν_i 表征。适当时,应该给出协方差。
- 3) B类分量应该用 u_j^2 表征。可以认为 u_j^2 是假设存在的相应方差的近似。可以像方差那样处理 u_j^2 , 并像标准差那样去处理 u_j 。适当时,也应该以同样方法给出协方差。
- 4) 合成不确定度应该采用常用的方差合成方法得到的数值表征。合成不确定度及其分量应该用“标准差”的形式表示。
- 5) 对于特殊应用,有必要用合成不确定度乘以一个因子获得总不确定度,并且应声明被乘因子。

测量不确定度评定和表示

1 范围

1.1 本标准规定了测量不确定度评定和表示的通用规则,适用于从生产车间到基础研究等很多领域的各种准确度水平的测量,包括:

- a) 生产过程中利用测量活动进行的质量控制和质量保证;
- b) 法律和法规中涉及的测量结果的符合性判定;
- c) 科学和工程领域的基础研究、应用研究和开发工作中的测量活动;
- d) 为溯源到国家测量标准,对测量标准和仪器进行的校准;
- e) 研制、保存国际和国家物理测量标准,包括标准物质,以及开展比对。

1.2 本标准主要涉及已有明确定义并可用唯一值表征的物理量的测量结果的不确定度表示方法。如果某个现象仅呈现为一系列值的分布或取决于一个或若干个参数,例如时间,则被测量就是包含其分布情况或相互关系的一组量值。

1.3 本标准也适用于实验、测量方法、复杂部件和系统的设计和理论分析中相关不确定度的评定与表示。因为测量结果及其不确定度可能只是概念上的,可能完全基于假设的数据,所以在本标准中按广义来理解术语“测量结果”。

1.4 本标准仅提供了测量不确定度评定和表示的通用规则,而不是详尽的技术规范文件。此外,本标准不讨论某个特定的测量结果的不确定度在其评定后如何用于不同的目的,例如:判断一个结果与另一些类似结果间是否相容;确定制造工艺中的容差极限;决定是否能安全地保证某个工作过程等。所以,可能有必要以此标准为基础制定一些专用标准,以解决特定测量领域的一些特殊问题或不确定度定量表示的各种用途。

注:在有些情况下可能会认为测量不确定度概念不完全适用,例如当确定某种测试方法的精密度时(例如,参考文献[5])。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的,凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 3358.1—2009 统计学词汇及符号 第1部分:一般统计术语与用于概率的术语(ISO 3534-1:2006^[7], IDT)

ISO/IEC 指南 99:2007 国际计量学词汇 基础和通用概念及相关术语(International vocabulary of metrology—Basic and general concepts and associated terms)(VIM)^[6]

3 术语和定义

ISO/IEC 指南 99:2007(VIM)界定的计量学术语和定义,GB/T 3358.1 界定的统计学术语和定义,以及下列术语和定义适用于本文件。

3.1

测量不确定度 measurement uncertainty; uncertainty of measurement

不确定度 uncertainty

利用可获得的信息,表征赋予被测量量值分散性的非负参数。

注 1: 测量不确定度包括由系统效应引起的分量, 如与修正量和测量标准所赋量值有关的分量及定义的不确定度。有时对估计的系统效应未作修正, 而是当作不确定度分量处理。

注 2: 此参数可以是诸如称为标准测量不确定度的标准差(或其特定倍数), 或是说明了包含概率的区间半宽度。

注 3: 测量不确定度一般由若干分量组成。其中一些分量可根据一系列测量值的统计分布, 按测量不确定度的 A 类评定进行评定, 并可用标准差表征。而另一些分量则可根据经验或其他信息所获得的概率密度函数, 按测量不确定度的 B 类评定进行评定, 也用标准差表征。

注 4: 通常, 对于一组给定的信息, 测量不确定度是相应于所赋予被测量的值的, 该值的改变将导致相应的不确定度的改变。

3.2

标准不确定度 **standard uncertainty**

以标准差表示的测量不确定度。

3.3

[标准不确定度的]A 类评定 **type A evaluation (of standard uncertainty)**

对在规定测量条件下测得的量值用统计分析的方法进行的测量不确定度分量的评定。

3.4

[标准不确定度的]B 类评定 **type B evaluation (of standard uncertainty)**

用不同于测量不确定度 A 类评定的方法对测量不确定度分量进行的评定。

示例:

评定基于以下信息:

- 权威机构发布的量值;
- 有证参考物质的量值;
- 校准证书;
- 仪器的漂移;
- 经检定的测量仪器的准确度等级;
- 根据人员经验推断的极限值等。

3.5

合成标准不确定度 **combined standard uncertainty**

由在一个测量模型中各输入量的标准测量不确定度获得的输出量的标准测量不确定度。

注: 如果测量模型中的输入量相关, 当计算合成标准不确定度时应考虑协方差。

3.6

扩展不确定度 **expanded uncertainty**

合成标准测量不确定度与一个大于 1 的数字因子的乘积。

注 1: 该因子取决于测量模型中输出量的概率分布类型及所选取的包含概率。

注 2: 本定义中术语“因子”是指包含因子。

注 3: 扩展测量不确定度在 INC-1(1980)建议书的第 5 段中被称为“总不确定度”。

3.7

包含因子 **coverage factor**

为获得扩展不确定度, 对合成标准不确定度所乘的大于 1 的数。

注: 包含因子通常用符号 k 表示。

4 标准不确定度的评定

4.1 测量模型

4.1.1 很多情况下, 被测量 Y 不能直接测得, 而是由 N 个其他量 X_1, X_2, \dots, X_N 通过函数关系 f 来

确定:

$$Y = f(X_1, X_2, \dots, X_N) \dots\dots\dots (1)$$

注 1: 为简洁,对物理量(被测量)和代表该量的可能观察到的随机变量(见 4.2.1),在本标准中使用同一个符号。当声明 X_i 具有特定的概率分布时,这个符号的含义就是后者;并假设该物理量本身可以用一个唯一值来表征(见 1.2)。

注 2: 在一列观测值中,第 k 个 X_i 的观测值用 $X_{i,k}$ 表示;因此,如果电阻器的电阻符号为 R ,则电阻的第 k 次观测值表示为 R_k 。

注 3: X_i 的估计值(严格说来是它的期望的估计值)用 x_i 表示。

示例:

如果一个随温度变化的电阻器两端的电压为 V ,在温度为 t_0 时的电阻为 R_0 ,电阻器的温度系数为 α ,在温度 t 时电阻器的损耗功率 P (被测量)取决于 V, R_0, α 和 t :

$$P = f(V, R_0, \alpha, t) = V^2 / \{R_0 [1 + \alpha(t - t_0)]\}$$

注: 测量损耗功率 P 的其他方法可能有不同的数学模型。

4.1.2 输出量 Y 的输入量 X_1, X_2, \dots, X_N 本身可看作被测量,其本身也可能取决于其他量,包括修正系统影响的修正值和修正因子,从而导致一个复杂的函数关系式,以至函数 f 不能明确地表示出来。甚至, f 可以用实验方法确定(见 5.1.4)或者只用数值方程给出。在本标准中函数 f 是按此广义来理解的,尤其当函数包含了对测量结果不确定度有显著影响的所有量,包括全部修正值和修正因子在内时更是如此。

因此,如果数据表明 f 没有将测量过程模型化至测量所要求的准确度,则应在 f 中增加输入量以弥补不足。这就可能需要引入一个输入量来反映对影响被测量的现象认识不足。在 4.1.1 的例子中,需要考虑将电阻器上已知的非均匀温度分布、电阻的非线性温度系数,或电阻与大气压力的关系作为附加的输入量。

注: 尽管如此,式(1)可以简单为 $Y = X_1 - X_2$ 。例如,将对同一量 X 进行两次测量的比较时,模型的表达式就如此简单。

4.1.3 一组输入量 X_1, X_2, \dots, X_N 可以分类为:

- 其值和不确定度是在当前的测量过程中直接测定的量。这些值和不确定度可从单次观测、重复观测或经验判断获得,并可能包含测量仪器读数的修正值,以及对环境温度、大气压力和湿度等影响量的修正值。
- 其值和不确定度是由外部来源获得引入测量过程的量,例如与已校准的测量标准,有证标准物质,从手册中查到的参考数据相关的量。

4.1.4 被测量 Y 的估计值用 y 表示,是用 N 个输入量 X_1, X_2, \dots, X_N 的估计值 x_1, x_2, \dots, x_N 代入式(1)得到。因此输出量的估计值 y ,即测量结果,由式(2)给出:

$$y = f(x_1, x_2, \dots, x_N) \dots\dots\dots (2)$$

注: 在某些情况下,估计值 y 可由下式得到:

$$y = \bar{Y} = \frac{1}{n} \sum_{k=1}^n Y_k = \frac{1}{n} \sum_{k=1}^n f(X_{1,k}, X_{2,k}, \dots, X_{N,k})$$

也就是说, y 是取 Y 的 n 次独立观测值 Y_k 的算术平均值(见 4.2.1),其每个观测值 Y_k 的不确定度相同,且每个 Y_k 都是根据同时获得的 N 个输入量 X_i 的一组完整观测值求得的。当 f 是输入量 X_1, X_2, \dots, X_N 的非线性函数时,用此式求平均值比用 $y = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N)$ 更具优越性:其中 $\bar{X}_i = \frac{1}{n} \sum_{k=1}^n X_{i,k}$, \bar{X}_i 是独立观测值 $X_{i,k}$ 的算术平均值,但如果 f 是 X_i 的线性函数时,则两种方法是一样的(见 F.2 和 F.4)。

4.1.5 输出量的估计值或测量结果 y 的标准差估计值称为合成标准不确定度,用 $u_c(y)$ 表示,它决定于每个输入量的估计值 x_i 的标准差估计值,称为标准不确定度,用 $u_c(x_i)$ 表示。

4.1.6 每个输入量的估计值 x_i 及其相关的标准不确定度 $u_c(x_i)$ 是由输入量 X_i 的可能值的分布获得的。这个概率分布可能是根据 X_i 的一系列观测值 $x_{i,k}$ 得到的频率分布,或者也可能是一种先验分布。

标准不确定度分量的 A 类评定根据的是频率分布,而 B 类评定根据的是先验分布。无论上述哪一种情况,分布都只是代表人们对它的认识程度的模型。

4.2 标准不确定度的 A 类评定

4.2.1 在大多数情况下,随机变量 q 的期望值 μ_q 的最佳估计值是 n 次独立观测值的算术平均值或均值 \bar{q} ,其中 n 个独立观测值 q_k 是在相同测量条件下得到的:

$$\bar{q} = \frac{1}{n} \sum_{k=1}^n q_k \quad \dots\dots\dots(3)$$

因此,当从 n 次独立重复观测值 $x_{i,k}$ 估计输入量 X_i 时,由式(3)得到的样本平均值 \bar{X}_i 被用作式(2)中的输入量的估计值 x_i ,也就是说 $x_i = \bar{X}_i$,以便求得测量结果 y 。没有进行重复观测获得的输入量的估计值应由其他方法获得,如 4.1.3 中第二类评定方法中所给的那样。

4.2.2 由于影响量的随机变化或随机效应,每次独立观测值 q_k 亦不同。观测值的实验方差即 q 的概率分布的方差 σ^2 的估计值为:

$$s^2(q_k) = \frac{1}{n-1} \sum_{j=1}^n (q_j - \bar{q})^2 \quad \dots\dots\dots(4)$$

这个方差的估计值及其正平方根 $s(q_k)$ 表征了被观测值 q_k 的变异性,或更确切地说,表征了它们在平均值 \bar{q} 附近的分散性。 $s(q_k)$ 称为实验标准差。

4.2.3 平均值方差 $\sigma^2(\bar{q}) = \sigma^2/n$ 的最佳估计值由式(5)确定:

$$s^2(\bar{q}) = \frac{s^2(q_k)}{n} \quad \dots\dots\dots(5)$$

平均值的实验标准差 $s(\bar{q})$ 等于平均值的实验方差 $s^2(\bar{q})$ 的正平方根, $s(\bar{q})$ 和 $s^2(\bar{q})$ 均表征了 \bar{q} 与 q 的期望值 μ_q 的接近程度,两者都可用作 \bar{q} 的不确定度的度量。

因此,对于由 n 次独立重复观测值 $X_{i,k}$ 确定输入量 X_i 时其估计值 $x_i = \bar{X}_i$ 的标准不确定度 $u(x_i)$ 为 $u(x_i) = s(\bar{X}_i)$,其中 $s^2(\bar{X}_i)$ 根据式(5)计算得到。为方便起见, $u^2(x_i) = s^2(\bar{X}_i)$ 及 $u(x_i) = s(\bar{X}_i)$ 有时分别被称为 A 类方差和 A 类标准不确定度。

注 1: 观测次数 n 应该充分多,以使 \bar{q} 成为随机变量 q 的期望值 μ_q 的可靠估计值,并使 $s^2(\bar{q})$ 成为算术平均值 \bar{q} 的方差 $\sigma^2(\bar{q}) = \sigma^2/n$ (见 4.3.2,注)的可靠估计值。在设定包含区间(见 6.2.2)时应考虑 $s^2(\bar{q})$ 和 $\sigma^2(\bar{q})$ 之间的区别。在这种情况下,如果 q 的概率分布是正态分布(见 4.3.4),则这一区别可通过 t 分布来考虑(见 E.3.2)。

注 2: 虽然方差 $s^2(\bar{q})$ 是更为基本的量,在实际工作中用标准差 $s(\bar{q})$ 更为方便,因为它具有与 q 相同的量纲并且比方差更易于理解的值。

4.2.4 为了在统计控制下更好地表征测量过程,可以使用合成或合并方差估计值 s_p^2 (或合并实验标准差 s_p)。在这种情况下,如果被测量 q 的值是依据 n 次独立观测值确定的,那么用 s_p^2/n 作为算术平均值 \bar{q} 的实验方差的估计值,将优于使用 $s^2(q_k)/n$,标准不确定度 $u = s_p/\sqrt{n}$ (见 F.3.6 的注)。

4.2.5 通常输入量 X_i 的估计值 x_i 由实验数据用最小二乘法拟合一条曲线得到,表征曲线拟合参数和曲线上任何预期点的估计方差以及标准不确定度可以用已知的统计程序评定(见 F.3 和参考文献[8])。

4.2.6 在简单情况下, $u(x_i)$ (见 E.3) 的自由度 ν_i 等于 $n-1$,此时, $x_i = \bar{X}_i$ 和 $u(x_i) = s(\bar{X}_i)$ 由 n 次独立观测值按 4.2.1 和 4.2.3 方法计算得到,当以文件形式说明不确定度分量的 A 类评定时,通常应该给出其自由度。

4.2.7 如果一个输入量在多次观测中存在相关的随机效应,例如与时间相关,则用 4.2.1 和 4.2.3 的方法计算的平均值和平均值的实验标准差可能是这些统计量的不恰当估计。在这种情况下,应该采用专门为相关的一系列随机变化测量的数据处理设计的统计方法来分析观测值。

注: 这些专门的方法被用于处理频率标准的测量。但在对其他量的测量中,当测量从短期测量过渡到长期测量时,

仍假设为不相关的随机变化可能就不合理了,这种测量也需要用专门的方法来处理(见参考文献[9],例如对阿仑方差的详细讨论)。

4.2.8 4.2.1 到 4.2.7 对标准不确定度的 A 类评定并没有详尽论述;还有许多情况,有些甚至是很复杂的情况,可以用统计方法处理。校准设计方案的应用是一个很好的例子,通常用最小二乘法评定未知值的实物量具(如量块和标准砝码)与已知值的参考标准比较结果的短期和长期随机变化引入的不确定度。在这种相对简单的测量情形下,不确定度分量通常通过数据的统计分析来评定,这些数据是在一套测量程序中将其有关的量设计许多不同值来获得,这就是所谓的方差分析(见 F.5)。

注:在溯源链较低端的校准,通常假设参考标准的值是准确已知的,因为是经过国家或原级标准实验室校准过的,校准结果的不确定度可以从表征测量的合并实验标准差评定得到的单一 A 类标准不确定度。

4.3 标准不确定度的 B 类评定

4.3.1 当输入量 X_i 的估计值 x_i 不由重复观测得到时,估计方差 $u^2(x_i)$ 或标准不确定度 $u(x_i)$ 可根据 X_i 可能变化的全部有关信息的判断来评定,信息来源可能包括:

- 以前的测量数据;
- 对有关材料和仪器特性的经验或了解;
- 生产厂提供的技术说明书;
- 校准证书或其他证书提供的数据;
- 手册给出的参考数据的不确定度。

为方便起见,用这种方法估计的 $u^2(x_i)$ 和 $u(x_i)$ 有时分别称为 B 类方差和 B 类标准不确定度。

注:当 x_i 由一个先验分布得到时,其方差恰当的写法是 $u^2(X_i)$,但为简单起见,本标准中都写成 $u^2(x_i)$ 和 $u(x_i)$ 。

4.3.2 在 B 类标准不确定度的评定中,合理使用所得到的信息需要经验以及知识,这是一种可以在实践中掌握的技能。应该认识到,标准不确定度的 B 类评定可以与 A 类评定一样可靠,特别是当 A 类评定基于的统计独立观测次数较少时。

注:如果 4.2.3 的注 1 中 q 的概率分布为正态分布,那么 $s(\bar{q})$ 的标准差与 $\sigma(\bar{q})$ 之比 $\sigma[s(\bar{q})]/\sigma(\bar{q})$ 近似为 $[2(n-1)]^{-1/2}$ 。因此,取 $\sigma[s(\bar{q})]$ 为 $s(\bar{q})$ 的不确定度,当观测次数 $n=10$ 时, $s(\bar{q})$ 的相对不确定度为 24%,当 $n=50$ 时则为 10%(其他值在附录 C 的表 C.1 中给出)。

4.3.3 如果估计值 x_i 取自制造厂的技术说明书、校准证书、手册或其他来源,并且明确给出了其不确定度是标准差的若干倍,则标准不确定度 $u(x_i)$ 可取为给定值除以该倍数所得之商,而估计方差 $u^2(x_i)$ 是商的平方。

示例:

校准证书声明标称值为 1 kg 的标准砝码的质量 $m_s = 1\,000.000\,32$ g,并说明该值的不确定度按 3 倍标准差计为 0.24 mg。则该标准砝码的标准不确定度 $u(m_s) = (0.24\text{ mg})/3 = 0.08$ mg。相应的相对标准不确定度 $u(m_s)/m_s$ 为 80×10^{-9} (见 5.1.6)。其估计方差 $u^2(m_s) = (0.08\text{ mg})^2 = 6.4 \times 10^{-9} \text{ g}^2$ 。

注:在很多情况下,能获得给定不确定度分量的信息很少甚至没有。一般来说这一点并不重要,因为根据本标准来表示不确定度时,当计算测量结果的合成标准不确定度时所有的标准不确定度是用同样的方法处理的(见第 5 章)。

4.3.4 x_i 的不确定度不一定像 4.3.3 那样由标准差的倍数给出。可以将不确定度定义为具有 90%、95% 或 99% 包含概率的一个区间(见 6.2.2)。除另有说明外,可以假定不确定度是用正态分布来计算的,而 x_i 的标准不确定度可以通过将该不确定度除以对应正态分布的包含因子来得到。对应上述三个包含概率的因子分别为 1.645, 1.960, 2.576(见附录 E 的表 E.1)。

注:如果已经根据本标准推荐的方法给出不确定度,就不需要再作正态分布的假设,本标准在报告不确定度时强调的是要给出所用的包含因子(见 7.2.3)。

示例:

标称值为 10 Ω 的标准电阻器的校准证书中给出的校准结果为:阻值 R_s 在 23 $^{\circ}\text{C}$ 时为 $10.000\,74\ \Omega \pm 0.13\ \text{m}\Omega$,并说明

“给定的 0.13 mΩ 不确定度所对应的区间具有 99% 的包含概率”。则电阻器的标准不确定度 $u(R_S) = (0.13 \text{ m}\Omega) / 2.58 = 0.05 \text{ m}\Omega$, 相对标准不确定度 $u(R_S) / R_S = 5 \times 10^{-6}$ (见 5.1.6)。估计方差为 $u^2(R_S) = (0.05 \text{ m}\Omega)^2 = 2.5 \times 10^{-9} \Omega^2$ 。

4.3.5 考虑下面的情况, 根据得到的信息, 输入量 X_i 的值有 50% 的机会落在 a_- 到 a_+ 的区间内(换言之, X_i 的值落在这个区间的概率为 0.5 或 50%)。如果假设 X_i 的可能值近似为正态分布, 那么 X_i 的最佳估计值 x_i 可取为该区间的中点。进而, 设该区间的半宽度为 $a = (a_+ - a_-) / 2$ 表示, 则可以取 $u(x_i) = 1.48a$, 因为对一个数学期望为 μ , 标准差为 σ 的正态分布, 区间 $\mu \pm \sigma / 1.48$ 约包含分布的 50%。

示例:

一个机械师在测定零件的尺寸时, 估计其长度以 0.5 的概率落在 10.07 mm 到 10.15 mm 之间, 并报告长度 $l = (10.11 \pm 0.04) \text{ mm}$, 这说明 $\pm 0.04 \text{ mm}$ 对应 50% 包含概率的区间。因此 $a = 0.04 \text{ mm}$, 假设 l 的可能值服从正态分布, 长度 l 的标准不确定度 $u(l) = 1.48 \times 0.04 \text{ mm} \approx 0.06 \text{ mm}$, 方差的估计值 $u^2(l) = (1.48 \times 0.04 \text{ mm})^2 = 3.5 \times 10^{-3} \text{ mm}^2$ 。

4.3.6 考虑与 4.3.5 类似的情况, 根据获得的信息声明: X_i 的值约以 2/3 的可能性落在 a_- 到 a_+ 的区间内(换言之, X_i 的值落在这个区间的概率为 0.67)。则可合理地取 $u(x_i) = a$, 因为对一个数学期望为 μ , 标准差为 σ 的正态分布, 区间 $\mu \pm \sigma$ 约包含分布的 68.3%。

注: 若用对应于概率 $p = 2/3$ 的实际正态偏离 0.967 42, 即写成 $u(x_i) = a / 0.967 42 = 1.033a$, 则给出的 $u(x_i)$ 值比上述所取的值要合理得多。

4.3.7 在另一些情况下, 可能估计 X_i 的极限(上限和下限), 尤其当声明“为实用起见, X_i 落在 a_- 到 a_+ 区间内的概率等于 1, 而 X_i 落在这个区间外的概率为零”。如果对 X_i 的可能值落在该区间内的情况缺乏更多的认识, 只能假设 X_i 是按等概率落在该区间的任何地方[可能值服从均匀分布或矩形分布——见 4.4.5 和图 2a]。则 X_i 的数学期望或期望值 x_i 是该区间的中点, $x_i = (a_- + a_+) / 2$, 其方差为:

$$u^2(x_i) = (a_+ - a_-)^2 / 12 \quad \dots\dots\dots (6)$$

如果用 $2a$ 表示上限与下限之差 $a_+ - a_-$, 则式(6)变为:

$$u^2(x_i) = a^2 / 3 \quad \dots\dots\dots (7)$$

注: 当用这种方法得到的不确定度分量对测量结果的不确定度贡献较大时, 为谨慎起见, 需要获得其他数据来作进一步评定。

示例 1:

手册给出了纯铜在 20 °C 时的线膨胀系数 $\alpha_{20}(\text{Cu})$ 为 $16.52 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$, 且该值的误差不超过 $0.40 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ 。基于该有限的信息, 假设 $\alpha_{20}(\text{Cu})$ 以等概率落在 $16.12 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ 到 $16.92 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ 区间, 并且不太可能落在该范围之外是合理的。区间半宽度为 $a = 0.40 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$, 由式(7), $\alpha_{20}(\text{Cu})$ 的对称矩形分布的方差 $u^2(\alpha_{20}) = (0.40 \times 10^{-6} \text{ }^\circ\text{C}^{-1})^2 / 3 = 53.3 \times 10^{-15} \text{ }^\circ\text{C}^{-2}$, 标准不确定度 $u(\alpha_{20}) = (0.40 \times 10^{-6} \text{ }^\circ\text{C}^{-1}) / \sqrt{3} = 0.23 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ 。

示例 2:

一台数字电压表的制造厂说明书写明: “仪器校准后的 2 年内, 1 V 量程电压准确度是: 读数 $\times (14 \times 10^{-6}) +$ 量程 $\times (2 \times 10^{-6})$ ”。假定仪器校准后使用了 20 个月, 在 1 V 量程测量电压 V , V 的一组独立重复观测值的算术平均值 $\bar{V} = 0.928 571 \text{ V}$, 其 A 类标准不确定度为 $u(\bar{V}) = 12 \mu\text{V}$ 。B 类标准不确定度可依据制造厂说明书, 假定该引用的准确度数据对 \bar{V} 的修正值 $\Delta\bar{V}$ 是对称边界, $\Delta\bar{V}$ 的期望为零, 并等概率地落在边界内的任意地方。 $\Delta\bar{V}$ 可能值的对称矩形分布的半宽度 a 为: $a = (14 \times 10^{-6}) \times (0.928 571 \text{ V}) + (2 \times 10^{-6}) \times (1 \text{ V}) = 15 \mu\text{V}$, 并由式(7)得到 $u^2(\Delta\bar{V}) = \left(\frac{a}{\sqrt{3}}\right)^2 = 75 \mu\text{V}^2$, $u(\Delta\bar{V}) = 8.7 \mu\text{V}$ 。为简单起见, 被测量 V 的估计值使用同样的符号 V 来表示, $V = \bar{V} + \Delta\bar{V} = 0.928 571 \text{ V}$ 。此估计值的合成标准不确定度由 \bar{V} 的 A 类标准不确定度 $12 \mu\text{V}$ 与 $\Delta\bar{V}$ 的 B 类标准不确定度 $8.7 \mu\text{V}$ 合成得到。标准不确定度通用的合成方法可参见第 5 章, 此例子在 5.1.5 中讨论。

4.3.8 在 4.3.7 中, 输入量 X_i 的上限 a_+ 和下限 a_- 相对于其最佳估计值 x_i 不一定是对称的; 更准确地说: 如果下限 $a_- = x_i - b_-$, 而上限 $a_+ = x_i + b_+$, 其中 $b_- \neq b_+$ 。在这种情况下 x_i (假设 x_i 为 X_i 的期望值) 不在 a_- 到 a_+ 区间的中心, X_i 的概率分布在区间内不会是均匀的。但是, 可能没有足够的信息来选择合适的分布; 不同的模型会导致不同的方差表达式。在缺乏信息时, 最简单的近似式为:

$$u^2(x_i) = \frac{(b_+ + b_-)^2}{12} = \frac{(a_+ - a_-)^2}{12} \dots\dots\dots(8)$$

这就是全宽度为 $b_+ + b_-$ 的矩形分布的方差(在 D.2.4.4 和 E.5.3 中也讨论了不对称分布)。

示例:

如果在 4.3.7 的示例 1 中,手册中给出的线膨胀系数值为 $\alpha_{20}(\text{Cu}) = 16.52 \times 10^{-6} \text{ } ^\circ\text{C}^{-1}$,并声明“最小可能值是 $16.40 \times 10^{-6} \text{ } ^\circ\text{C}^{-1}$,最大可能值是 $16.92 \times 10^{-6} \text{ } ^\circ\text{C}^{-1}$ ”,则 $b_- = 0.12 \times 10^{-6} \text{ } ^\circ\text{C}^{-1}$, $b_+ = 0.40 \times 10^{-6} \text{ } ^\circ\text{C}^{-1}$,并由式(8)得 $u(\alpha_{20}) = 0.15 \times 10^{-6} \text{ } ^\circ\text{C}^{-1}$ 。

注 1:在许多实际测量中,界限是不对称的。此时可适当对估计值 x_i 加一个修正值 $(b_+ - b_-)/2$,则 X_i 新的估计值 x'_i 就落在界限的中心位置 $x'_i = (a_- + a_+)/2$ 。这样就可归到 4.3.7 的情况了,其新值 $b'_+ = b'_- = (b_+ + b_-)/2 = (a_+ - a_-)/2 = a$ 。

注 2:根据最大熵原理,在不对称情况下的概率密度函数可以表示为:

$$p(X_i) = A \exp[-\lambda(X_i - x_i)]$$

式中:

$$A = [b_- \exp(\lambda b_-) + b_+ \exp(-\lambda b_+)]^{-1}$$

$$\lambda = \{\exp[\lambda(b_- + b_+)] - 1\} / \{b_- \exp[\lambda(b_- + b_+)] + b_+\}$$

其方差为 $u^2(x_i) = b_+ b_- - (b_+ - b_-)/\lambda$;在 $b_+ > b_-$ 时, $\lambda > 0$,在 $b_+ < b_-$ 时 $\lambda < 0$ 。

4.3.9 在 4.3.7 中,因为对 X_i 的可能值落在估计界限 a_- 到 a_+ 内的情况缺乏具体的了解,所以只能假设 X_i 在界限内等概率地取任意值,而在界限外取值的概率为零。在概率分布中这种阶跃式不连续函数通常是不符合实际的。很多情况下,更现实的是靠近界限的值比靠近中点的值要少。因此用斜边相等的对称梯形分布(一个等腰梯形)来代替对称矩形分布更为合理,梯形分布的底部宽 $a_+ - a_- = 2a$,顶部宽为 $2a\beta$,其中 $0 \leq \beta \leq 1$ 。当 $\beta \rightarrow 1$ 时,此梯形分布接近于 4.3.7 的矩形分布,而当 $\beta = 0$ 时就成为三角分布[见 4.4.6 和图 2b)]。假设 X_i 服从梯形分布, X_i 的期望值 $x_i = (a_+ + a_-)/2$,其方差为:

$$u^2(x_i) = a^2(1 + \beta^2)/6 \dots\dots\dots(9a)$$

当 $\beta = 0$ 时,变为三角分布,其方差为:

$$u^2(x_i) = a^2/6 \dots\dots\dots(9b)$$

注 1:对一个期望为 μ ,标准差为 σ 的正态分布来说,区间 $\mu \pm 3\sigma$ 大约包含了分布的 99.73%。因而,如果上、下限 a_+ 和 a_- 是按 99.73% 而不是按 100% 的界限定义的,并且 X_i 可被认为近似正态分布而不是如在 4.3.7 中那样对 X_i 在界限内的情况完全不了解,此时 $u^2(x_i) = a^2/9$ 。可对比下,半宽度为 a 的对称矩形分布的方差为 $a^2/3$ [式(7)],半宽度为 a 的对称三角分布的方差为 $a^2/6$ [式(9b)]。这三种分布的方差的数量级很接近,但从判断它们所需的信息量来说,区别是很大的。

注 2:梯形分布等于两个矩形分布的卷积^[10],一个矩形分布的半宽度 a_1 等于梯形分布的平均半宽度, $a_1 = a(1 + \beta)/2$,另一个矩形分布的半宽度 a_2 等于梯形分布的一个三角部分的平均宽度, $a_2 = a(1 - \beta)/2$ 。该梯形分布的方差 $u^2 = a_1^2/3 + a_2^2/3$ 。该卷积分布可以被看成一个宽度为 $2a_1$ 的矩形分布,而该分布自身又具有宽度为 $2a_2$ 的矩形分布的不确定度,卷积分布就是这样模拟了一个并不准确知道输入量边界的分布。但即使 a_2 大到 a_1 的 30%, u 超过 $a_1/\sqrt{3}$ 的可能性仍不到 5%。

4.3.10 避免“重复计算”不确定度分量非常重要。如果产生自某种特殊效应的不确定度分量是用 B 类评定方法获得的,仅当该效应对观测值的变异性没有贡献时,它才应该在计算测量结果的合成标准不确定度时作为独立的不确定度分量。如果该效应对观测值的变异性有贡献,则其导致的不确定度已经包含在观测值的统计分析里了。

4.3.11 在 4.3.3 到 4.3.9 中对 B 类标准不确定度评定的讨论仅仅是指导性的。如在 3.4.1 和 3.4.2 中所强调的,不确定度的评定应该尽最大可能以定量的数据为根据。

4.4 评定标准不确定度的图解说明

4.4.1 图 1 表示了由一个输入量 X_i 可能测得值的未知分布,或用重复观测的方法采样获得的 X_i 的概率分布得到 X_i 的估计值及其不确定度。

4.4.2 在图 1a)中假设输入量 X_i 是温度 t , 其未知分布假设为正态分布, 期望值 $\mu_t = 100\text{ }^\circ\text{C}$, 标准差 $\sigma = 1.5\text{ }^\circ\text{C}$, 其概率密度函数为:

$$p(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu_t}{\sigma}\right)^2\right]$$

注: 概率密度函数 $p(z)$ 的定义要求 $\int p(z)dz = 1$ 成立。

4.4.3 图 1b)显示从温度 t 的 $n=20$ 次重复观测值 t_k 得到的直方图, 假设从图 1a) 的分布中随机采样 t 值。为获得直方图, 将 20 个观测值或采样值按 $1\text{ }^\circ\text{C}$ 间隔分组(当然, 对数据进行统计分析并不需要绘制直方图), 观测值见表 1。

表 1 温度 t 按 $1\text{ }^\circ\text{C}$ 间隔分组的 20 次重复观测值

间隔 $t_1 \leq t \leq t_2$		温度 $t/^\circ\text{C}$
$t_1/^\circ\text{C}$	$t_2/^\circ\text{C}$	
94.5	95.5	—
95.5	96.5	—
96.5	97.5	96.90
97.5	98.5	98.18, 98.25
98.5	99.5	98.61, 99.03, 99.49
99.5	100.5	99.56, 99.74, 99.89, 100.07, 100.33, 100.42
100.5	101.5	100.68, 100.95, 101.11, 101.20
101.5	102.5	101.57, 101.84, 102.36
102.5	103.5	102.72
103.5	104.5	—
104.5	105.5	—

$n=20$ 次观测值的算术平均值 \bar{t} 可根据式(3)计算得到, $\bar{t} = 100.145\text{ }^\circ\text{C} \approx 100.14\text{ }^\circ\text{C}$, 根据数据得到的 \bar{t} 是 t 的期望值 μ_t 的最佳估计值。实验标准差 $s(t_k)$ 由式(4)计算得到, $s(t_k) = 1.489\text{ }^\circ\text{C} \approx 1.49\text{ }^\circ\text{C}$, 平均值的实验标准差 $s(\bar{t})$ 由式(5)计算得到, 它就是平均值 \bar{t} 的标准不确定度 $u(\bar{t})$, $u(\bar{t}) = s(\bar{t}) = s(t_k) / \sqrt{20} = 0.333\text{ }^\circ\text{C} \approx 0.33\text{ }^\circ\text{C}$ (为了作进一步计算, 可能要保留全部数字)。

注: 表 1 所列数据仅用来说明问题, 不要理解为是实际情况。

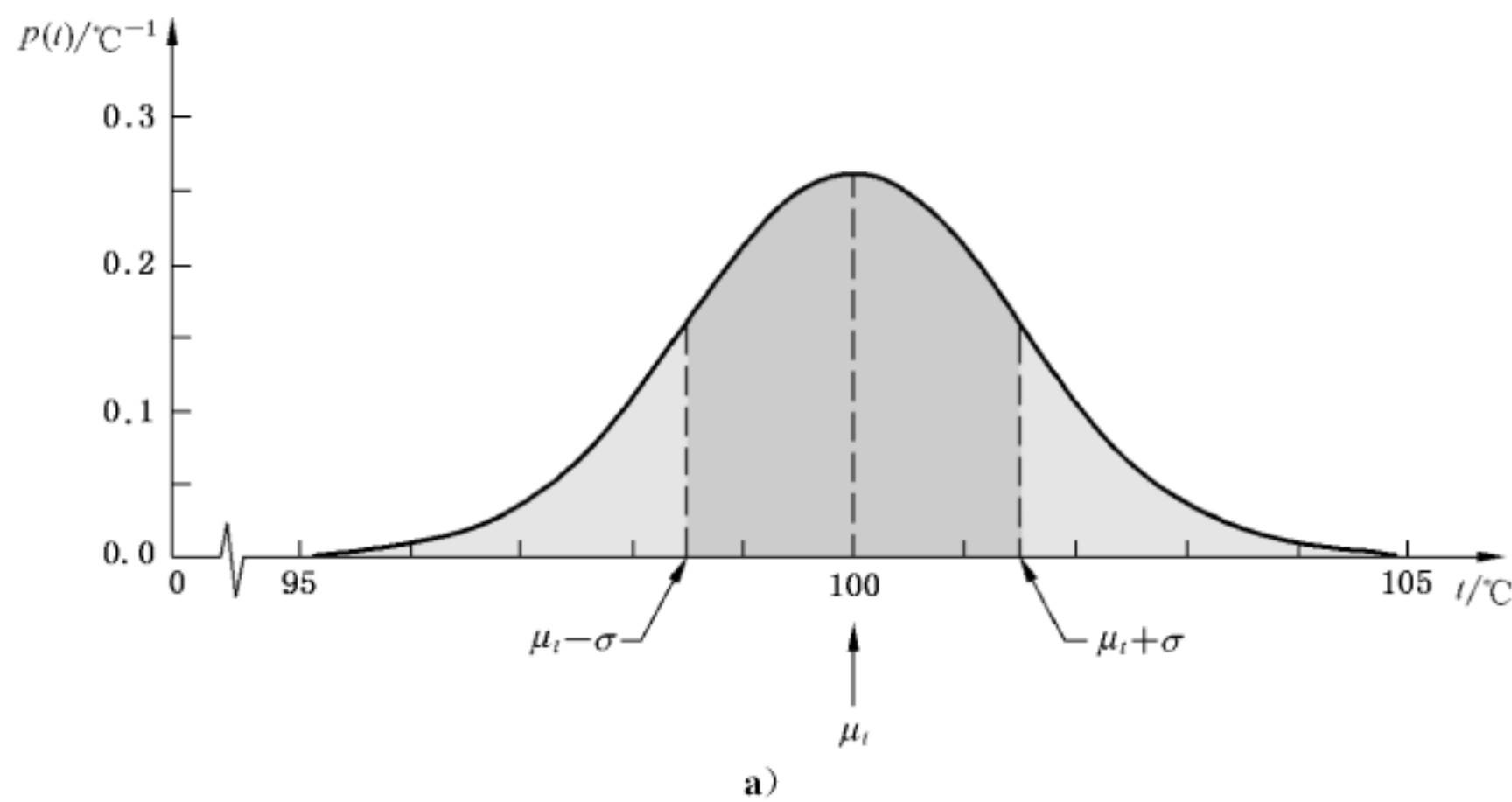


图 1 由重复观测评定输入量的标准不确定度示意图

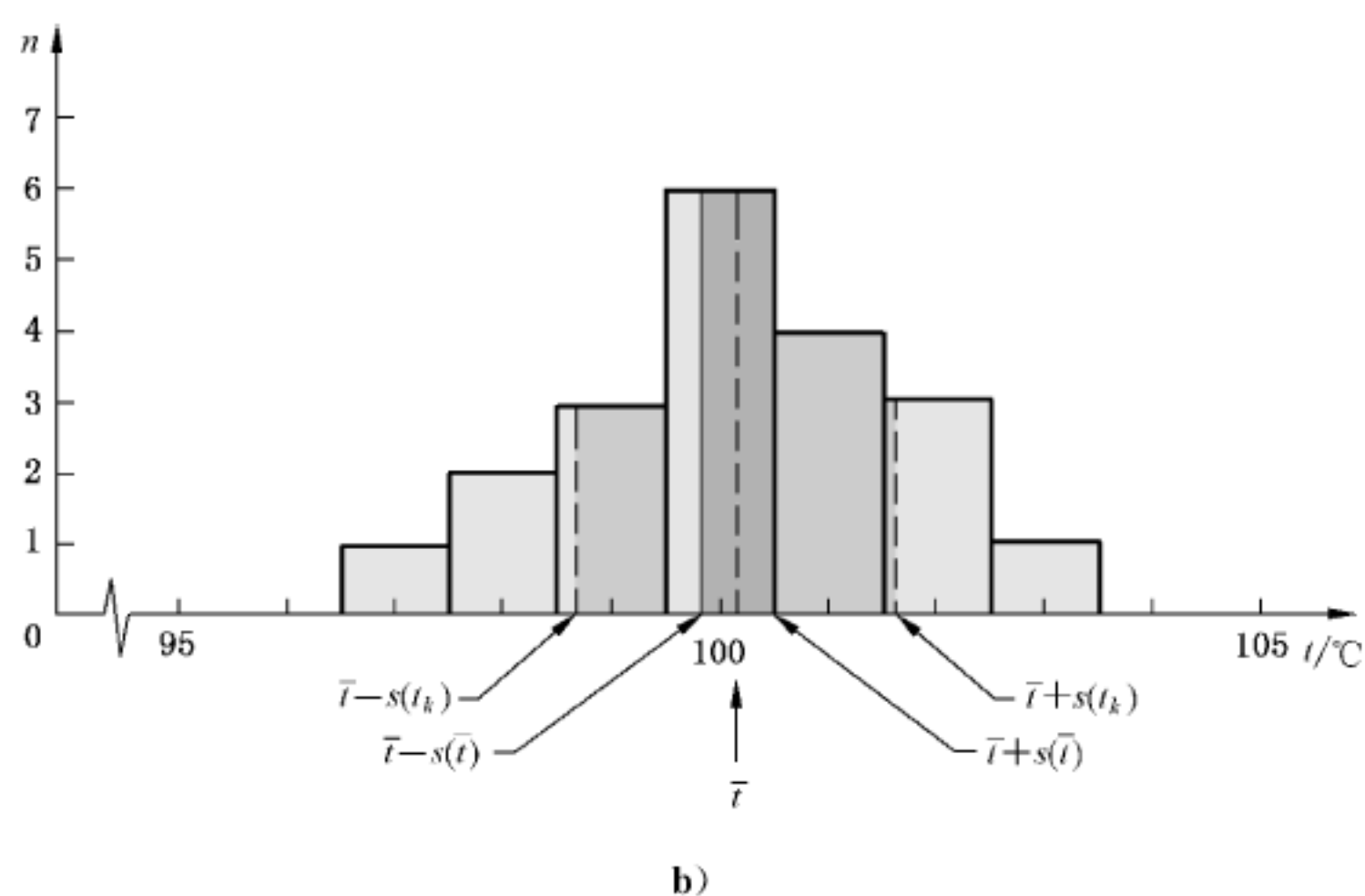


图 1 (续)

4.4.4 图 2 所示为由输入量 X_i 的可能值的先验分布或基于全部可获得信息得到的 X_i 的概率分布获得的 X_i 的估计值以及评定的该估计值的不确定度。对这两种情况,输入量仍假设为温度 t 。

4.4.5 在图 2a) 中,如果对于输入量 t 几乎没有可利用的信息,唯一可假设的是 t 可用一个对称矩形的先验概率分布来描述,其下限 $a_- = 96\text{ }^\circ\text{C}$,上限 $a_+ = 104\text{ }^\circ\text{C}$,因而半宽度 $a = (a_+ - a_-)/2 = 4\text{ }^\circ\text{C}$ (见 4.3.7)。则 t 的概率密度函数为:

$$p(t) = 1/(2a), \text{ 当 } a_- \leq t \leq a_+ \text{ 时};$$

$$p(t) = 0, \text{ 当 } t \text{ 取其他值时}.$$

正如在 4.3.7 指出的, t 的最佳估计值是它的期望值 $\mu_t = (a_+ + a_-)/2 = 100\text{ }^\circ\text{C}$ 。这个估计值的标准不确定度为 $u(\mu_t) = a/\sqrt{3} \approx 2.3\text{ }^\circ\text{C}$, 这是由式(7)得出的。

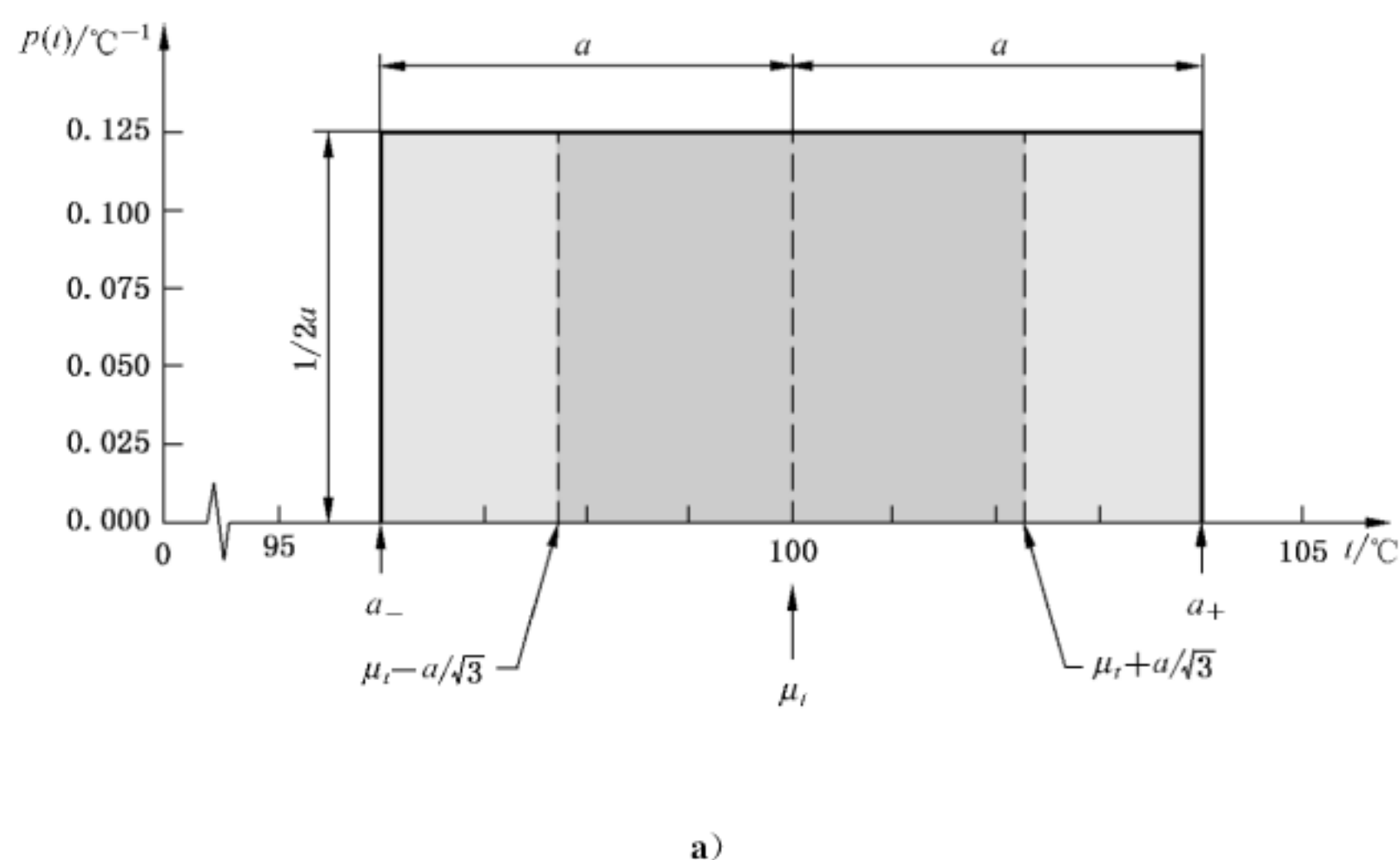


图 2 由先验分布评定输入量的标准不确定度示意图

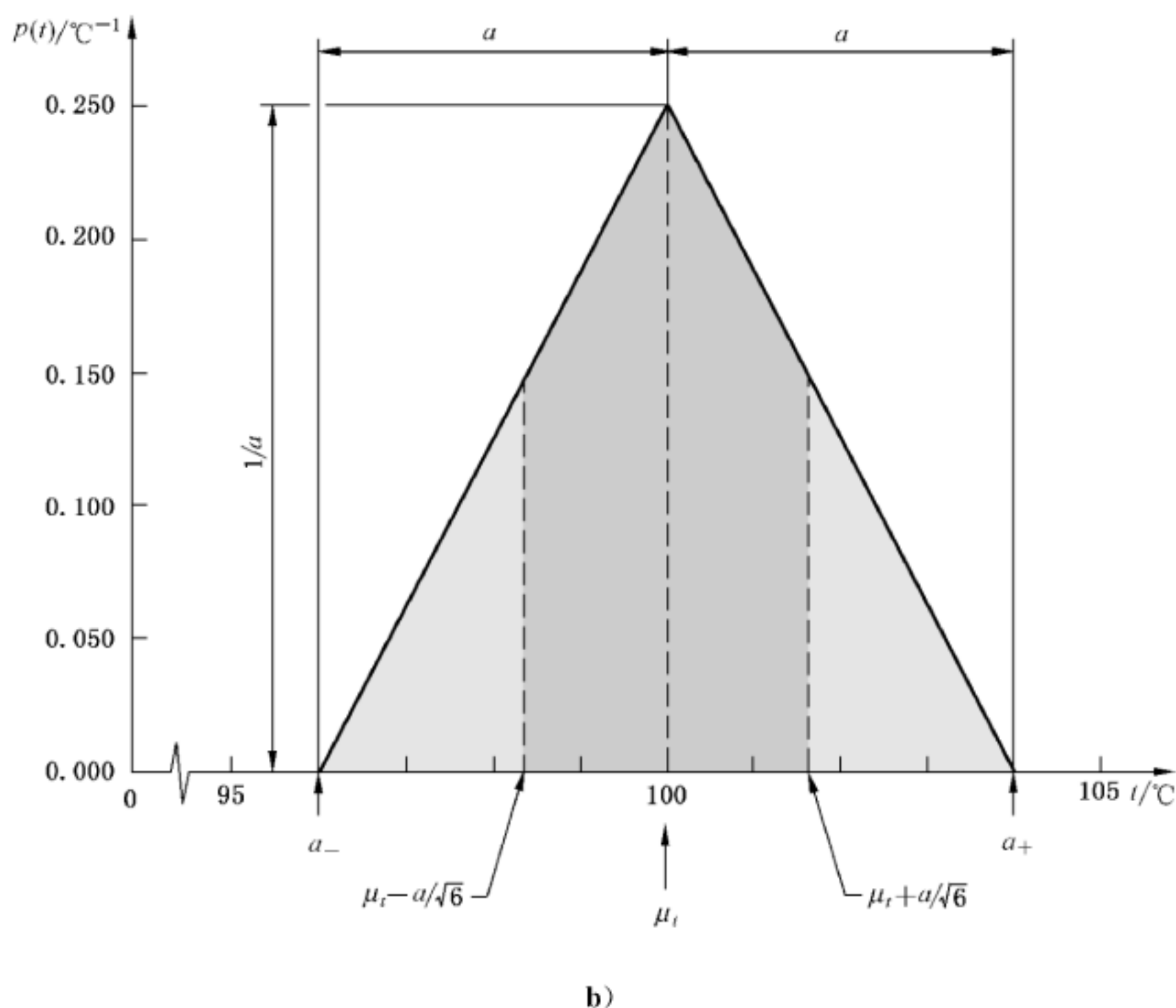


图 2 (续)

4.4.6 在图 2b)中,如果关于输入量 t 还有些可利用信息,并且 t 可用对称三角形先验概率分布描述,其下限仍为 $a_- = 96\text{ }^\circ\text{C}$,上限为 $a_+ = 104\text{ }^\circ\text{C}$,其半宽度如同 4.4.5 中一样为 $a = (a_+ - a_-)/2 = 4\text{ }^\circ\text{C}$ (见 4.3.9)。则 t 的概率密度函数为:

$$\begin{aligned}
 p(t) &= (t - a_-)/a^2, \text{ 当 } a_- \leq t \leq (a_+ + a_-)/2 \text{ 时;} \\
 p(t) &= (a_+ - t)/a^2, \text{ 当 } (a_+ + a_-)/2 \leq t \leq a_+ \text{ 时;} \\
 p(t) &= 0, \text{ 当 } t \text{ 取其他值时。}
 \end{aligned}$$

如在 4.3.9 中所指出的, t 的期望值 $\mu_t = (a_+ + a_-)/2 = 100\text{ }^\circ\text{C}$,这是根据 B.3.1 得出的。估计值的标准不确定度 $u(\mu_t) = a/\sqrt{6} \approx 1.6\text{ }^\circ\text{C}$,这是根据 B.3.2 得出的[见式 9b)]。

上述 $u(\mu_t) = 1.6\text{ }^\circ\text{C}$,可与 4.4.5 中宽度同样为 $8\text{ }^\circ\text{C}$ 的矩形分布 $u(\mu_t) = 2.3\text{ }^\circ\text{C}$ 相比较;图 1a)的正态分布的 $\sigma = 1.5\text{ }^\circ\text{C}$,由 -2.58σ 到 $+2.58\sigma$ 的宽度近似为 $8\text{ }^\circ\text{C}$,包含了分布的 99%;其不确定度 $u(\bar{t}) = 0.33\text{ }^\circ\text{C}$,该值是在 4.4.3 中假设从同一正态分布随机选取 20 次观测值得到的。

关于评定不确定度分量的实用指南可参见附录 F。

5 合成标准不确定度的计算

5.1 不相关的输入量

本节讨论的是所有输入量均为相互独立(B.3.7)的情况。当两个或两个以上输入量相关时,即互不独立或相关的情况在 5.2 中讨论。

5.1.1 当 y 是被测量 Y 的估计值因而也就是测量结果时, y 的标准不确定度由输入量的估计值 x_1, x_2, \dots, x_N (见 4.1)的标准不确定度的适当合成得到。估计值 y 的合成标准不确定度用 $u_c(y)$ 表示。

注:与 4.3.1 的注中的理由相同,在所有情况下都用符号 $u_c(y)$ 和 $u_c^2(y)$ 。

5.1.2 合成标准不确定度 $u_c(y)$ 是合成方差 $u_c^2(y)$ 的正平方根,由式(10)给出:

$$u_c^2(y) = \sum_{i=1}^N \left[\frac{\partial f}{\partial x_i} \right]^2 u^2(x_i) \quad \dots\dots\dots(10)$$

其中 f 是式(1)给出的函数。每个 $u(x_i)$ 是按 4.2(A类评定)或 4.3(B类评定)所评定的标准不确定度。合成标准不确定度 $u_c(y)$ 是一个估计标准差,表征可合理赋予被测量 Y 的值的分散性。式(10)及其相关输入量的对应式(13)都是基于 $Y=f(X_1, X_2, \dots, X_N)$ 的泰勒级数的一阶近似,在本标准中把它称为不确定度传播律(见 C.3.1 和 C.3.2)。

注:当 f 为明显非线性时,在 $u_c^2(y)$ 的表达式(10)中应包括泰勒级数展开中的高阶项。当每个 X_i 都是正态分布时,加到式(10)中去的下一高阶的主要项为:

$$\sum_{i=1}^N \sum_{j=1}^N \left[\frac{1}{2} \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 + \frac{\partial f}{\partial x_i} \frac{\partial^3 f}{\partial x_i \partial x_j^2} \right] u^2(x_i) u^2(x_j)$$

在 F.1 中给出了需要考虑较高阶项的示例。

5.1.3 偏导数 $\partial f/\partial x_i$ 是在 $X_i=x_i$ 时 $\partial f/\partial X_i$ 的值(见下面注 1),常把这些导数称为灵敏系数,它描述输出量的估计值 y 如何随输入量的估计值 x_1, x_2, \dots, x_N 的值的 变化而变化。尤其,输入量的估计值 x_i 的微小变化 Δx_i 引起的 y 的变化可用 $(\Delta y)_i = (\partial y/\partial x_i)(\Delta x_i)$ 表示。如果这个变化是由估计值 x_i 的标准不确定度所引起的,则 y 的相应变化为 $(\partial f/\partial x_i)u(x_i)$ 。因此合成方差 $u_c^2(y)$ 可以看作各项之和,每一项代表了由每一个输入量的估计值 x_i 相关联的估计方差产生的输出量的估计值 y 的估计方差。由此,建议把式(10)写成:

$$u_c^2(y) = \sum_{i=1}^N [c_i u(x_i)]^2 \equiv \sum_{i=1}^N u_i^2(y) \quad \dots\dots\dots(11a)$$

式中:

$$c_i \equiv \partial f/\partial x_i, \quad u_i(y) \equiv |c_i| u(x_i) \quad \dots\dots\dots(11b)$$

注 1:严格说来,偏导数 $\partial f/\partial x_i$ 应是 $\partial f/\partial X_i$ 在 X_i 的期望值上的值,然而在实际工作中,偏导数是按下式估计的:

$$\frac{\partial f}{\partial x_i} = \left. \frac{\partial f}{\partial X_i} \right|_{x_1, x_2, \dots, x_N}$$

注 2:合成标准不确定度 $u_c(y)$ 可以用下式代替式(11a)中的 $c_i u(x_i)$,进行数值运算得到:

$$Z_i = \frac{1}{2} \{ f[x_1, \dots, x_i + u(x_i), \dots, x_N] - f[x_1, \dots, x_i - u(x_i), \dots, x_N] \}$$

即 $u_i(y)$ 是用计算由于 x_i 变化 $\pm u(x_i)$ 而引起的 y 变化来评定的。于是可取 $u_i(y)$ 为 $|Z_i|$ 的值,而相应的灵敏系数 c_i 为 $Z_i/u(x_i)$ 。

注 3:式中符号“ \equiv ”表示“恒等于”。

示例:

用 4.1.1 的例子,为简单起见对量和它的估计值采用同样的符号,

$$c_1 \equiv \partial P/\partial V = 2V/\{R_0[1+\alpha(t-t_0)]\} = 2P/V$$

$$c_2 \equiv \partial P/\partial R_0 = -V^2/\{R_0^2[1+\alpha(t-t_0)]\} = -P/R_0$$

$$c_3 \equiv \partial P/\partial \alpha = -V^2(t-t_0)/\{R_0[1+\alpha(t-t_0)]^2\} = -P(t-t_0)/[1+\alpha(t-t_0)]$$

$$c_4 \equiv \partial P/\partial t = -V^2\alpha/\{R_0[1+\alpha(t-t_0)]^2\} = -P\alpha/[1+\alpha(t-t_0)]$$

以及:

$$\begin{aligned} u^2(P) &= \left[\frac{\partial P}{\partial V} \right]^2 u^2(V) + \left[\frac{\partial P}{\partial R_0} \right]^2 u^2(R_0) + \left[\frac{\partial P}{\partial \alpha} \right]^2 u^2(\alpha) + \left[\frac{\partial P}{\partial t} \right]^2 u^2(t) \\ &= [c_1 u(V)]^2 + [c_2 u(R_0)]^2 + [c_3 u(\alpha)]^2 + [c_4 u(t)]^2 \\ &= u_1^2(P) + u_2^2(P) + u_3^2(P) + u_4^2(P) \end{aligned}$$

5.1.4 灵敏系数 $\partial f/\partial x_i$ 有时不是通过函数 f 计算得到,而是用实验确定:采用变化一个特定的 X_i ,而将其余输入量保持不变时,测量出由此引起的 Y 的变化。这种情况下,对函数 f (或当仅有几个灵敏系数是如此确定时,对函数 f 的一部分)的了解就相应地缩小到基于测得的灵敏系数进行的经验性的一阶泰勒级数展开式。

5.1.5 如果被测量 Y 的式(1)在输入量 X_i 的名义值 $X_{i,0}$ 处作一阶展开时(通常这是一种恰当的近似), $Y=Y_0+c_1\delta_1+c_2\delta_2+\dots+c_N\delta_N$, 其中 $Y_0=f(X_{1,0}, X_{2,0}, \dots, X_{N,0})$, c_i 是导数 $\partial f/\partial X_i$ 在 $X_i=X_{i,0}$ 处的值, $\delta_i=X_i-X_{i,0}$ 。为了分析不确定度, 通常将输入量 X_i 变换到 δ_i , 以使被测量近似为其变量的线性函数(见 C.3.1)。

示例:

4.3.7 的例 2 中, 被侧量 V 的估计值是 $V=\bar{V}+\Delta V$, 其中 $\bar{V}=0.928\ 571\ \text{V}$, $u(\bar{V})=12\ \mu\text{V}$, 附加修正值 $\Delta\bar{V}=0$, $u(\Delta\bar{V})=8.7\ \mu\text{V}$ 。由于 $\partial V/\partial\bar{V}=1$, $\partial V/\partial(\Delta V)=1$, 则 \bar{V} 的合成方差为

$$u_c^2(V) = u^2(\bar{V}) + u^2(\Delta\bar{V}) = (12\ \mu\text{V})^2 + (8.7\ \mu\text{V})^2 = 219 \times 10^{-12}\ \text{V}^2$$

所以合成标准不确定度为 $u_c(V)=15\ \mu\text{V}$, 其相应的相对合成标准不确定度 $u_c(V)/V$ 为 16×10^{-6} (见 5.1.6)。这就是被测量为其依据的量的线性函数, 且系数 $c_i=+1$ 情况的例子。如果 $Y=c_1X_1+c_2X_2+\dots+c_NX_N$, 并且常数 $c_i=+1$ 或 -1 , 则由式(10)得出 $u_c^2(y) = \sum_{i=1}^N u^2(x_i)$ 。

5.1.6 如果 $Y=cX_1^{p_1} X_2^{p_2} \dots X_N^{p_N}$, 指数 p_i 是已知正数或负数并且其不确定度可忽略, 则合成方差式(10)可以表示为:

$$[u_c(y)/y]^2 = \sum_{i=1}^N [p_i u(x_i)/x_i]^2 \dots\dots\dots(12)$$

这与式(11a)具有同样的形式, 只不过其合成方差 $u_c^2(y)$ 用相对合成方差 $[u_c(y)/y]^2$ 表示, 与每一个输入量的估计值相关的方差估计值 $u^2(x_i)$ 表示为相对方差的估计值 $[u(x_i)/x_i]^2$ 。[相对合成标准不确定度为 $u_c(y)/|y|$, 各个输入量的估计值的相对标准不确定度为 $u(x_i)/|x_i|$, 其中 $|y| \neq 0, |x_i| \neq 0$ 。]

注 1: 当 Y 具有这种形式时, 可设 $X_i=X_{i,0}(1+\delta_i)$, 方便地将 Y 变换成变量的线性函数(见 5.1.5), 然后得到以下近

似关系式的结果, $(Y-Y_0)/Y_0 = \sum_{i=1}^N p_i \delta_i$ 。另一方面, 对数变换 $Z=\ln Y$ 和 $W_i=\ln X_i$ 可以使新的变量完全线性化为 $Z=\ln c + \sum_{i=1}^N p_i W_i$ 。

注 2: 如果每个 p_i 只是 $+1$ 或 -1 , 式(12)就变成 $[u_c(y)/y]^2 = \sum_{i=1}^N [u(x_i)/x_i]^2$, 表明了在这种情况下, 估计值 y 的相对合成方差等于输入量的估计值 x_i 的相对方差估计值之和。

5.2 相关的输入量

5.2.1 式(10)及其导出式, 如式(11a)和式(12), 仅适用于输入量 X_i 间独立或不相关的情况(随机变量, 不是假设为不变的物理量——见 4.1.1, 注 1)。如果某些 X_i 是明显相关时就应考虑其相关性。

5.2.2 当输入量相关时, 测量结果的合成方差 $u_c^2(y)$ 的表达式为:

$$u_c^2(y) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) = \sum_{i=1}^N \left[\frac{\partial f}{\partial x_i} \right]^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \dots\dots\dots(13)$$

其中 x_i 和 x_j 是 X_i 和 X_j 的估计值, 且 $u(x_i, x_j)=u(x_j, x_i)$ 是 x_i 和 x_j 的协方差的估计值。 x_i 和 x_j 之间的相关程度用估计的相关系数表征(B.3.6):

$$r(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i)u(x_j)} \dots\dots\dots(14)$$

其中 $r(x_i, x_j) = r(x_j, x_i)$, 且 $-1 \leq r(x_i, x_j) \leq +1$, 如果估计值 x_i 和 x_j 是相互独立的, 则 $r(x_i, x_j)=0$, 即一个值的变化不会意味着另一个值有预期的变化(更多信息见 B.3.6, B.3.7)。

相关系数这个术语比协方差易于解释, 式(13)中的协方差项可以写成:

$$2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i)u(x_j)r(x_i, x_j) \dots\dots\dots(15)$$

再用式(11b)代入后,使式(13)变成:

$$u_c^2(y) = \sum_{i=1}^N c_i^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N c_i c_j u(x_i) u(x_j) r(x_i, x_j) \dots\dots\dots (16)$$

注 1: 非常特殊情况下,所有的输入量的估计值都相关,且相关系数 $r(x_i, x_j) = +1$,此时式(16)可简化为:

$$u_c^2(y) = \left[\sum_{i=1}^N c_i u(x_i) \right]^2 = \left[\sum_{i=1}^N \frac{\partial f}{\partial x_i} u(x_i) \right]^2$$

此时合成标准不确定度 $u_c(y)$ 是由每个输入量的估计值 x_i 的标准不确定度产生的输出量的估计值 y 的变化之线性和(见 5.1.3)。[不应该将这种线性和与误差传播律相混淆,虽然形式相同,但标准不确定度不是误差(见 C.3.2)。]

示例:

有 10 个电阻器,每个电阻器的标称值为 $R_i = 1 \text{ k}\Omega$,用 $1 \text{ k}\Omega$ 的标准电阻器 R_s 校准,比对的不确定度可忽略,标准电阻器的不确定度由校准证书给出为 $u(R_s) = 100 \text{ m}\Omega$ 。将这些电阻器用电阻可忽略的导线串连起来,得到标称值为 $10 \text{ k}\Omega$ 的参考电阻 R_{ref} 。因此 $R_{\text{ref}} = f(R_i) = \sum_{i=1}^{10} R_i$ 。由于对每对电阻器来说 $r(x_i, x_j) = r(R_i, R_j) = 1$ (见 D.1.2.3 中例 2),本注中的公式是适用的。由于每个电阻器 $\partial f / \partial x_i = \partial R_{\text{ref}} / \partial R_i = 1$, $u(x_i) = u(R_i) = u(R_s)$ (见 D.1.2.3 例 2),由公式得到 R_{ref} 的合成标准不确定度 $u_c(R_{\text{ref}}) = \sum_{i=1}^{10} u(R_s) = (10 \times 100) \text{ m}\Omega = 1 \text{ }\Omega$ 。由式(10)获得的结果 $u_c(R_{\text{ref}}) = \left[\sum_{i=1}^{10} u^2(R_s) \right]^{1/2} = 0.32 \text{ }\Omega$ 是不正确的。因为没有考虑 10 个电阻的被校值是相关的。

注 2: 估计方差 $u^2(x_i)$ 和估计协方差 $u(x_i, x_j)$ 可以考虑为一个元素为 u_{ij} 的协方差矩阵的元素。矩阵的对角线元素 u_{ii} 是方差 $u^2(x_i)$,而非对角线的元素 $u_{ij} (i \neq j)$ 是协方差 $u(x_i, x_j) = u(x_j, x_i)$ 。如果两个输入量的估计值是不相关的,则其相关的协方差及协方差阵中相应的元素 u_{ij} 和 u_{ji} 为 0。如果输入量的估计值均不相关,则非对角线的元素均为 0,协方差矩阵就是对角矩阵(见 B.3.5)。

注 3: 为进行数值运算,式(16)可写成:

$$u_c^2(y) = \sum_{i=1}^N \sum_{j=1}^N Z_i Z_j r(x_i, x_j)$$

其中 Z_i 在 5.1.3 的注 2 中已说明。

注 4: 如果在 5.1.6 中考虑的特殊形式的 X_i 是相关的,则应在式(12)的右边增加以下项:

$$2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N [p_i u(x_i) / x_i] [p_j u(x_j) / x_j] r(x_i, x_j)$$

5.2.3 考虑到算术平均值 \bar{q} 和 \bar{r} 分别为两个随机变量 q 和 r 的期望值 μ_q 和 μ_r 的估计值,并在同一测量条件下对 q 和 r 进行 n 次独立同时重复观测所得值计算得到(见 B.2.15), \bar{q} 和 \bar{r} 的协方差用式(17)估计(见 B.3.4):

$$s(\bar{q}, \bar{r}) = \frac{1}{n(n-1)} \sum_{k=1}^n (q_k - \bar{q})(r_k - \bar{r}) \dots\dots\dots (17)$$

其中 q_k 和 r_k 是量 q 和 r 的单次观测值, \bar{q} 和 \bar{r} 是由观测值根据式(3)计算得到的。如果实际上观测值是不相关的,则计算得到的协方差可预期接近零。

由此可见,两个相关输入量 X_i 和 X_j 是由独立重复同时观测的数据对平均值 \bar{X}_i 和 \bar{X}_j 评定的,而 X_i 和 X_j 的估计协方差 $u(x_i, x_j) = s(\bar{X}_i, \bar{X}_j)$,其中 $s(\bar{X}_i, \bar{X}_j)$ 是根据式(17)计算得到的。应用式(17)是协方差的 A 类评定。 \bar{X}_i 和 \bar{X}_j 的相关系数估计值由式(14)得到:

$$r(x_i, x_j) = r(\bar{X}_i, \bar{X}_j) = s(\bar{X}_i, \bar{X}_j) / [s(\bar{X}_i) s(\bar{X}_j)]$$

注: 在 F.2 和 F.3 中介绍了有必要按式(17)计算协方差的例子。

5.2.4 如果测量两个输入量时采用了同一台测量仪器、实物测量标准或参考数据,并且它们具有较大的标准不确定度,则两个输入量间就可能具有显著的相关性。例如用某一温度计来确定输入量 X_i 的估计值的温度修正值,并用同一温度计来确定输入量 X_j 的估计值的温度修正值,这两个输入量就可能显著

相关。然而,如果在这个例子中的 X_i 和 X_j 重新定义为不相关的量,而把确定温度计校准曲线的量作为附加的输入量,该附加输入量具有独立的标准不确定度,则此时 X_i 和 X_j 的相关性就去除了(更多信息见 D.1.2.3 和 D.1.2.4)。

5.2.5 如果输入量之间存在显著相关性时,则不能忽略。如果改变相关输入量的方法可行的话,应该采用实验方法评定相关方差,或者使用可获得的有关量的相关变异性信息(协方差的 B 类评定)进行评定。当估计由常有的影响量如环境温度、大气压力和湿度引起的输入量间的相关程度时,就特别需要经验和常识(见 4.3.1 和 4.3.2)。幸而在很多情况下,这些影响量之间的相关性很小,足以假设被影响的输入量是不相关的。然而,如果不能假设为不相关,则如 5.2.4 所指出的那样,把常有的影响量作为独立输入量引入,被影响的输入量本身之间的相关性就可避免了。

6 扩展不确定度的确定

6.1 扩展不确定度的由来

6.1.1 根据本标准所依据的测量不确定度声明工作组的建议书 INC-1(1980)(见本标准的引言)及 CIPM 批准并再次确认的建议书 1(CI-1981)和建议书 1(CI-1986)(见 A.2 和 A.3),提倡用合成标准不确定度 $u_c(y)$ 作为定量表示测量结果的不确定度的参数。实际上,在其第二次建议中,CIPM 已经要求所有参加 CIPM 及其咨询委员会举办的国际比对及其他工作的参加者在给出结果时使用合成标准不确定度 $u_c(y)$ 。

6.1.2 尽管 $u_c(y)$ 可以广泛用于表示测量结果的不确定度,但在某些商业、工业和法规的应用中以及涉及健康和安全时,常常有必要提供不确定度量,也就是给出测量结果值的区间,并期望该区间包含了能合理赋予被测量值分布的大部分。工作组承认了此需求的存在,并反映在建议书 INC-1(1980)的第 5 条中,在 CIPM 的建议书 1(CI-1986)中也有反映。

6.2 扩展不确定度

6.2.1 为满足提供一个如 6.1.2 所指的区间的要求而对不确定度的另一种度量称为扩展不确定度,并用 U 表示。扩展不确定度 U 由合成标准不确定度 $u_c(y)$ 乘以包含因子 k 得到:

$$U = k u_c(y) \dots\dots\dots (18)$$

由此,测量结果可方便地表示成 $Y = y \pm U$,意思是被测量 Y 的最佳估计值为 y ,由 $y - U$ 到 $y + U$ 是一个区间,可期望该区间包含了能合理赋予的 Y 值的分布的大部分。这样一个区间也可表示成 $y - U \leq Y \leq y + U$ 。

6.2.2 置信区间和置信水平这两个术语在统计学中具有专门的定义,本标准中对应这两个术语,使用“包含区间”和“包含概率”。其仅适用于满足某些条件的由 U 定义的区间,这些条件包括对 $u_c(y)$ 有贡献的所有不确定度分量都是由 A 类评定获得的。确切地说, U 可被理解为确定了测量结果的一个区间,该区间以较大的概率 p 包含了测量结果及其合成标准不确定度表征的概率分布的大部分, p 为区间的“包含概率”或“置信水平”。

6.2.3 可行时,应评估和声明由 U 确定的区间相关的包含概率 p 。应该认识到,用常数乘 $u_c(y)$ 并不提供新的信息,而是以不同的形式表示前面已获取的信息。然而,也应该认识到:在大多数情况下包含概率 p (尤其在 p 的值接近 1 时)是相当不确定的,不仅是因为对用 y 和 $u_c(y)$ 表征的概率分布的了解有限(尤其在边缘部分),而且因为 $u_c(y)$ 本身的不确定性(见 6.3.2,6.3.3 和附录 E,尤其 E.6.6)。

注:当不确定度用 $u_c(y)$ 和 U 表示时,建议分别按 7.2.2 和 7.2.4 声明测量结果。

6.3 包含因子的选择

6.3.1 包含因子 k 的值是根据 $y - U$ 到 $y + U$ 区间所要求的包含概率而选择的。一般 k 在 2~3 范围

内,但在某些应用领域 k 可能超出这个范围。应根据测量结果的使用的需要选取合适的 k 值,这可能需要对测量结果的使用具有丰富的经验和渊博的知识。

注:有时发现,系统影响的已知修正量 b 并未被用于修正报告的测量结果,而是通过放大结果的“不确定度”来将这种影响考虑在内,应该尽可能避免这样做。只有在非常特殊的情况下,虽然存在已知的明显的系统影响,仍不对测量结果进行修正(见 D.2.4.5 特殊情况及如何处理)。测量结果不确定度的评定不应该与对某些参数规定的安全极限相混淆。

6.3.2 理想情况下,能够选择包含因子 k 的特定值, k 值提供的区间 $Y=y\pm U=y\pm ku_c(y)$ 对应于特定的包含概率 p (例如 95%或 99%);同样,对于给定 k 值,也能明确地声明该区间有关的包含概率。然而在实践中不易做到,因为它需要对由测量结果 y 及其合成不确定度 $u_c(y)$ 表征的概率分布有全面的了解。虽然这些参数是非常重要的,但为建立具有准确已知的包含概率的区间,仅仅依据这些参数还不充分。

6.3.3 建议书 INC-1(1980)中并没有规定如何建立 k 和 p 的关系。这个问题在附录 E 中论述,在 E.4 中给出的建议是采用近似的方法,并在 E.6.4 中作了扼要总结。在 E.6.6 中讨论了一种较简便的方法,适用于当 y 和 $u_c(y)$ 所表征的概率分布近似为正态分布,且 $u_c(y)$ 的有效自由度较大时的测量情况。实际经常发生这种情况,可设 $k=2$,所形成的区间具有的包含概率约为 95%,设 $k=3$ 所形成的区间具有的包含概率约为 99%。

注:在 E.4 中给出了估计 $u_c(y)$ 的有效自由度的方法。附录 E 中的表 E.2 可用于帮助判断这种方法是否适用于特定的测量(见 E.6.6)。

7 不确定度的报告与表示

7.1 通用指南

7.1.1 通常,当人们要将测量等级升级时,就需要关于如何获得测量结果及其不确定度的更详细的信息。然而,无论在哪个等级上,包括市场中的贸易和监管活动、工业中的工程工作、处于溯源链中较低端的校准实验室、工业研究和开发、学术研究、工业部门的最高标准和校准中心及国家计量院和 BIPM,对重新评定测量所需的信息都应该让需要它的人获得。主要的差别在于:在溯源链的较低端,大部分必需的信息是以发布的校准和测试系统的报告、校准规范、校准证书、使用手册、国际标准、国家标准和地方法规等形式给出的。

7.1.2 当测量的细节,包括如何评定测量结果的不确定度,是通过引用发布的文件来提供时,如常见的用证书方式报告校准结果时,要求这些文件是最近的,使文件与实际使用的测量程序相一致。

7.1.3 工业和商业活动中每天都在进行大量的测量,但并没有明确地报告不确定度。然而,很多测量所用的器具是需要定期经过周期校准或依法检定的。如果知道仪器是符合其规范或符合现行标准文件的要求,其示值的不确定度可以由其规范或标准文件推断。

7.1.4 虽然实际中,可以根据测量结果的预期用途确定随测量结果提供哪些必要的信息,但对报告测量结果的基本原则是不变的:当报告测量结果及其不确定度时,宁可提供过多信息而不要太少。如应该包括:

- a) 清楚地描述由实验观测值和输入数据计算测量结果及其不确定度的方法;
- b) 将所有的不确定度分量列出,并将各分量如何评定的方法形成文件;
- c) 详细描述数据分析的方法,使其每个重要步骤易于遵照执行,必要时报告中结果的计算能独立重复;
- d) 给出在分析中所用的所有修正值和常数以及来源。

检查上述清单是否完备的方法就是自问,“我是否以足够清晰的方式提供了充分的信息,确保将来如果得到新的信息或数据,可以更新我的结果吗?”

7.2 专用指南

7.2.1 当用合成标准不确定度 $u_c(y)$ 报告测量结果的不确定度时,应当:

- 给出被测量 Y 的定义的充分描述;
- 给出被测量 Y 的估计值 y 及其合成标准不确定度 $u_c(y)$; y 和 $u_c(y)$ 都应当给出单位;
- 适当时,给出相对标准不确定度 $u_c(y)/|y|$, $|y| \neq 0$;
- 给出在 7.2.7 中所述的信息或包含此信息的公开文件。

如果认为对测量结果的用户有用,例如,有助于以后计算包含因子,或有助于对测量的理解,还可以指出:

- 估计的有效自由度 ν_{eff} (见 E.4);
- A 类和 B 类合成标准不确定度 $u_{cA}(y)$ 和 $u_{cB}(y)$ 及其估计的有效自由度 ν_{effA} 和 ν_{effB} (见 E.4.1 的注 3)。

7.2.2 当不确定度用 $u_c(y)$ 度量时,为避免误解,说明测量结果时最好用下列四种方法之一(假设被报告的量值是标称值为 100 g 的标准砝码,如果 u_c 在报告结果的文件中已有定义,则为简单起见下面括号中的文字可以忽略):

- “ $m_s = 100.021\ 47\ \text{g}$, (其合成标准不确定度) $u_c = 0.35\ \text{mg}$ ”;
- “ $m_s = 100.021\ 47(35)\ \text{g}$, 其中括号中的数是(合成标准不确定度) u_c 的数值, u_c 与测量结果的末位对齐”;
- “ $m_s = 100.021\ 47(0.000\ 35)\ \text{g}$, 括号中的数是(合成标准不确定度) u_c 的数值,以所给出结果的单位表示”;
- “ $m_s = (100.021\ 47 \pm 0.000\ 35)\ \text{g}$, 其中在“ \pm ”号后的数是(合成标准不确定度) u_c 的数值,不是一个包含区间”。

注:应避免使用士号的形式,因它已被传统地用于表示高包含概率的区间,因此可能会与扩展不确定度相混淆(见 7.2.4)。此外,虽然在方法 4)中加了防止误解的说明,但写成 $Y = y \pm u_c(y)$ 仍然可能会被误解。尤其如果偶尔将防止误解的说明遗漏了,就可能误解为 $k=1$ 的扩展不确定度,区间 $y - u_c(y) \leq Y \leq y + u_c(y)$ 具有确定的包含概率 p ,也就是说与正态分布有关(见 E.1.3)。如同在 6.3.2 和附录 E 中所指,用这种方式解释 $u_c(y)$ 通常是很难判别的。

7.2.3 当用扩展不确定度 $U = k u_c(y)$ 报告测量结果的不确定度时,应当做到:

- 给出被测量 Y 的定义的充分描述;
- 说明测量结果为 $Y = y \pm U$, 并给出 y 和 U 的单位;
- 适当时,给出相对扩展不确定度 $U/|y|$, $|y| \neq 0$;
- 给出获得 U 时所用的 k 值[或为用户方便,同时给出 k 和 $u_c(y)$];
- 给出与区间 $y \pm U$ 有关的近似的包含概率,并说明如何确定的;
- 给出在 7.2.7 中概述的信息或包含了上述内容的公开文件。

7.2.4 当不确定度用 U 度量时,为尽可能清晰,按如下示例说明测量结果的数值:

“ $m_s = (100.021\ 47 \pm 0.000\ 79)\ \text{g}$, 其中在“ \pm ”号后的数为扩展不确定度 $U = k u_c$ 的数值, U 是由 $u_c = 0.35\ \text{mg}$ 和 $k = 2.26$ 确定的, k 是基于 t 分布,自由度 $\nu = 9$ 得到的。确定区间估计具有约 95% 的包含概率”。

7.2.5 如果一次测量同时确定一个以上的被测量,也就是它提供两个或两个以上的输出量的估计值 y_i (见 F.2, F.3, F.4), 则除了给出 y_i 和 $u_c(y_i)$ 外,还要给出协方差矩阵的元素 $u(y_i, y_j)$ 或相关系数矩阵的元素 $r(y_i, y_j)$ (最好两个都给出)。

7.2.6 估计值 y 的数值及其标准不确定度 $u_c(y)$ 或扩展不确定度 U 不应该给出过多位数的数字。通常 $u_c(y)$ 和 U [以及输入量的估计值 x_i 的标准不确定度 $u(x_i)$] 最多为两位有效数字,虽然在某些情况

下,为了在随后计算中避免舍入误差而保留多余的位数。

在报告最终结果时,有时可能要将不确定度末位后面的数进位而不是舍去。例如, $u_c(y) = 10.47 \text{ m}\Omega$,可以进位到 $11 \text{ m}\Omega$ 。但一般的修约规则也可以用,如 $u(x_i) = 28.05 \text{ kHz}$,修约后为 28 kHz 。输入量和输出量的估计值应舍入到与其不确定度位数一致。例如,如果 $y = 10.057\ 62 \text{ }\Omega$,其 $u_c(y) = 27 \text{ m}\Omega$,则 y 应该进位到 $10.058 \text{ }\Omega$ 。如果相关系数的绝对值接近 1 时,则相关系数应给出三位有效数字。

7.2.7 当在报告中详细描述测量结果及其不确定度如何获得时,应当根据 7.1.4 的建议做到:

- a) 给出每个输入量的估计值 x_i 及其标准不确定度 $u(x_i)$,并说明它们是如何获得的;
- b) 对所有相关的输入量给出估计的协方差或估计的相关系数(最好两者都给),并说明获得的方法;
- c) 对每个输入量的估计值的标准不确定度给出自由度,并说明如何获得的;
- d) 给出函数关系 $Y = f(X_1, X_2, \dots, X_N)$ 。如果认为有用,给出其偏导数或灵敏系数 $\partial f / \partial x_i$ 。

但是,用实验方法确定的这种系数应当给出全部。

注:由于函数关系 f 可能非常复杂或者可能不存在显式而仅仅是一个计算程序,所以不是总能给出 f 及其导数的。函数 f 可能是一般描述或通过引用适当的参考文件给出,在这种情况下重要的是要搞清楚如何估计被测量 Y 的估计值 y 及如何得到其合成标准不确定度 $u_c(y)$ 。

8 不确定度评定与表示的程序概要

本标准中所述的评定和表示测量结果的不确定度的步骤归纳总结如下:

- 1) 将与被测量 Y 有关的输入量 X_i 与被测量 Y 间的关系用数学表达式表示出来: $Y = f(X_1, X_2, \dots, X_N)$,函数 f 应当包含每一个对测量结果的不确定度有显著影响的分量,包括所有的修正值和修正因子(见 4.1.1 和 4.1.2);
- 2) 既可基于一系列观测值的统计分析,也可用其他方法(见 4.1.3)确定输入量 X_i 的估计值 x_i ;
- 3) 评定每个输入量的估计值 x_i 的标准不确定度 $u(x_i)$ 。对由一系列观测值的统计分析获得的输入量的估计值,标准不确定度按 4.2 所述的方法评定(标准不确定度的 A 类评定)。对由其他方法得到的输入量的估计值,其不确定度 $u(x_i)$ 按 4.3 所述的方法评定(标准不确定度的 B 类评定);
- 4) 对任何相关的输入量,要评定他们的协方差(见 5.2);
- 5) 计算测量结果,即用步骤 2) 所得到的输入量 X_i 的估计值 x_i (见 4.1.4),通过函数关系式计算得到被测量 Y 的估计值 y ;
- 6) 按第 5 章所述,由输入量的估计值的标准不确定度和协方差来确定测量结果 y 的合成标准不确定度 $u_c(y)$ 。如果测量过程同时确定一个以上输出量,要计算它们的协方差(见 7.2.5, F.2, F.3 和 F.4);
- 7) 如果有必要给出扩展不确定度 U ,以便提供一个区间 $y - U$ 到 $y + U$,可期望该区间包含了能合理赋予被测量 Y 的值的分布的大部分,则将合成标准不确定度 $u_c(y)$ 乘以包含因子 k 得到 $U = k u_c(y)$, k 的值一般在 2~3 范围内。根据区间要求的包含概率选择 k (见 6.2, 6.3 和附录 E, 该附录讨论了 k 值的选择方法,由此产生一个具有接近规定的包含概率的区间);
- 8) 按 7.2.1 和 7.2.3 报告测量结果 y 及其合成标准不确定度 $u_c(y)$ 或扩展不确定 U ; 并且采用 7.2.2 和 7.2.4 中推荐的形式之一,并按第 7 章概述的方法说明如何获得 y , $u_c(y)$ 和 U 。

附 录 A

(资料性附录)

工作组和国际计量委员会(CIPM)的建议

A.1 建议书 INC-1(1980)

测量不确定度声明工作组(参见前言)是 BIPM 应 CIPM 的要求,在 1980 年 10 月成立的。工作组起草了一份详细的报告供 CIPM 参考,即建议书 INC-1(1980)^[2]。该建议书的内容见本标准引言 0.7。

A.2 建议书-1(CI-1981)

CIPM 在 1981 年 10 月召开的第 70 次大会上审议了测量不确定度声明工作组提交的报告,采纳了以下建议:

建议书-1(CI-1981)

实验不确定度的表示

CIPM 考虑到:

- 需要在计量学领域寻找一种公认的方式表述测量不确定度;
- 许多组织已经在这一领域付出了多年的努力;
- 1980 年从 BIPM 测量不确定度工作组的讨论,在找出公认的解决办法方面取得了可喜的进展。

认识到:

- 工作组的提议可以作为不确定度表示最终协议的基础。

建议:

- 广泛散发工作组的建议书;
- 今后 BIPM 努力在其主办的国际比对中应用该原则;
- 鼓励其他对此有兴趣的组织检验这些建议并将意见报告 BIPM;
- 在 2 年至 3 年后,BIPM 将报告这些建议的应用情况。

A.3 建议书-1(CI-1986)

CIPM 在 1986 年 10 月举行的第 75 次会议上进一步考虑了不确定度表示这一议题,通过如下建议:

建议书-1(CI-1986)

在 CIPM 主持下测量不确定度表示方面开展的工作

CIPM 考虑到:

- 不确定度工作组通过的建议书 INC-1(1980)和由 CIPM 通过的建议书-1(CI-1981)。
- 咨询委员会的某些成员可能寻求对该建议书进行解释,其目的是为了完成他们职权范围内的工作,尤其是为了进行国际比对。

认识到:

建议书 INC-1(1980)的第 5 条关系到特殊应用,尤其具有商业上的应用价值,现在正由国际标准化组织(ISO)的一个工作组进行考虑,该工作组是由 ISO、OIML 和 IEC 与 CIPM 共同组织的。

要求:

参加由 CIPM 及其咨询委员会主办的国际比对或其他工作的成员国在给出结果时应该用 INC-1(1980)建议书的第 4 条,并应给出用标准差表示的 A 类和 B 类不确定度的合成不确定度。

附录 B

(资料性附录)

“真”值、误差和不确定度及其他概念

本附录中 B.1 的主要内容来自 GUM:2008 的第 3 章“基本概念”;B.2 的主要内容来自 GUM:2008 的附录 D“‘真值’、误差、和不确定度”。

B.1 基本概念

B.1.1 测量

B.1.1.1 测量的目的是确定被测量的值,即待测的特定量的值,所以测量一开始就与恰当地规定被测量、测量方法和测量程序有关。

注:在本标准中不用术语“真值”;“被测量的值”(或量值)与“被测量的真值”(量的真值)被视为是等同的。

B.1.1.2 通常,量的测得值只是被测量值的近似或估计,因此只有当此估计值附有不确定度声明时,它才是完整的测量结果。

B.1.1.3 实际上,被测量的定义或规定的详细程度是随所要求的测量准确度而定的。针对所要求的准确度,被测量的定义应该足够完整,以便对与测量有关的所有实际目的而言,其值是唯一的。本标准中所用的“被测量的值”表示的就是这个意思。

示例:

若一根名义值为 1 m 长的钢棒需测至微米准确度,对该长度的定义应该包括确定长度时的温度和压力。例如,被测量应说明为:钢棒在 25.00 °C 和 101 325 Pa 时的长度(加上任何其他认为有必要的参数,如支撑钢棒的方式等)。然而,如果被测长度仅需测至毫米准确度,则对被测量的定义无需规定温度或压力或任何其他参数的值。

注:若被测量的定义不完整引入的测量不确定度分量足够大,在评定测量结果的不确定度时应包含此分量。

B.1.1.4 在很多情况下,测量结果是在重复性测量条件下得到的一系列观测值的基础上确定的。

B.1.1.5 重复观测中的变化是由于影响测量结果的影响量不能完全保持恒定而引起的。

B.1.1.6 将一组重复观测值变换成测量结果的测量数学模型是极为关键的,因为除观测值外,一般测量模型还包括各种不能确切知道的影响量。知识的缺乏也会是测量结果的不确定度的来源,如同重复观测的变化和数学模型本身带来的不确定度。

B.1.1.7 本标准仅把被测量作为标量(单一量)处理。当在同一个测量下同时测定一组相关的被测量时,要用矢量被测量和协方差矩阵代替标量被测量及其方差。本标准仅在举例中考虑了这种情况(见 F.2, F.3, F.4)。

B.1.2 误差、影响量和修正

B.1.2.1 通常,测量的不完善使测量结果存在误差。传统上认为误差有两类分量,即随机误差分量和系统误差分量。

注:误差是一个理想的概念,误差不可能准确获知。

B.1.2.2 随机误差一般是由影响量的不可预测性或其或在时间或空间上的随机变化所引起。这种变化量的影响以下称为随机影响,它引起被测量在重复观测中的变异性。测量结果的随机误差不能补偿,但通常可以通过增加观测次数来减小;其期望或期望值为零。

注 1:一系列观测值的算术平均值或均值的实验标准差(见 4.2.3)不是平均值的随机误差,虽然有些关于不确定度的出版物中是这样认为的。平均值的实验标准差是由随机影响引起的平均值的的不确定度的度量,而由于这些影响引起的平均值的误差的准确值是不可能知道的。

注2：术语“误差”和“不确定度”不是同义词，代表完全不同的概念，不应该混淆或误用。

B.1.2.3 系统误差与随机误差一样也是不可能被消除的，但通常也可以被减小。如果一个系统误差来源于测量结果的某个影响量中已识别的影响（以下称为系统影响），若这种影响可以定量给出，且其对测量所需的准确度而言比较大，则可用修正值或修正因子予以补偿。可以假设，修正后由系统影响引起的误差的期望或期望值为零。

注：用于补偿系统影响而加到测量结果上的修正值的不确定度不是系统误差，系统误差通常称为偏移。而修正值的不确定度是由于对修正值的认识不足引起的测量结果的不确定度的度量。对系统影响的不完全补偿所引起的误差是不可能准确获知的。

B.1.2.4 可以假定测量结果已经修正了所有已识别的重要的系统影响，并且已做了一切努力来识别这些影响。

示例：

用电压表测量高阻电阻器电位差（被测量）时，对电压表的阻抗进行修正可以减小电压表负载效应对测量结果的系统影响。然而电压表和电阻器的阻抗是用其他方法测量得到的，其测量本身也存在不确定度，而修正值又是使用该阻抗值估计的，因此电压表和电阻器阻抗的测量结果的不确定度成为电位差测量结果的不确定度分量。

注1：测量仪器和系统经常要使用测量标准和标准物质进行调整或校准，以便消除系统影响。然而，与这些测量标准和标准物质相关的不确定度应予以考虑。

注2：对已知的显著系统影响不进行修正的情况在 6.3.1 的注和附录 D.2.4.5 中讨论。

B.1.3 不确定度

B.1.3.1 测量结果的不确定度反映了对被测量的值缺乏精确的认识。对已识别的系统影响进行修正后的测量结果仍然只是被测量的估计值，因为还存在由随机影响引起的不确定度和由于对系统影响修正不完全而引入的不确定度。

注：测量结果（修正后）即使具有很大的不确定度，仍可能非常接近被测量的值（即误差可忽略）。因此，一个测量结果的不确定度不应与其剩余的未知误差相混淆。

B.1.3.2 实际上，测量中有许多不确定度的来源，包括：

- a) 被测量的定义不完整；
- b) 被测量定义的复现不完善；
- c) 取样的代表性不足——即所测量的样本可能不能代表定义的被测量；
- d) 对测量受环境条件的影响认识不足或对环境条件的测量不完善；
- e) 模拟指示仪器的人员读数偏移；
- f) 仪器分辨力或识别阈值的限制；
- g) 测量标准和标准物质的量值不准确；
- h) 从外部得到并在数据简约算法中使用的常数和其他参数的值不准确；
- i) 测量方法和程序中的近似和假设；
- j) 在看似相同条件下，被测量重复观测中的变异性。

这些来源不一定是相互独立的，a)到 i)中有些来源可能对来源 j)有贡献。当然，未识别的系统影响不可能在测量结果的不确定度评定中考虑到，但其对误差有贡献。

B.1.3.3 不确定度声明工作组起草的建议书 INC-1(1980)按不确定度的评定方法将不确定度分为两类：A类和B类（见 0.7, 3.3, 3.4），这种分类适用于“不确定度”，但“A类”和“B类”在用词上不能替代“随机”和“系统”。某个已知的系统影响的修正值的不确定度在某些情况下可能由A类评定方法得到；而在另一些情况下，可能由B类评定获得一个表征随机影响的不确定度。

注：在有些出版物中，将不确定度分量分为“随机”和“系统”分量，并分别将它们与由随机影响和已知系统影响引起的误差联系在一起。在常规应用中，这种不确定度分量的归类可能会不明确。例如，当前一个测量结果用作后

一测量的输入数据时,前一个测量的“随机”不确定度分量就变成后一个测量的“系统”不确定度分量。按不确定度分量的评定方法分类而不按分量自身性质分类避免了这一含混。同时,并不妨碍将已按两种不同方法评定的各分量归纳成组以便使用。

B.1.3.4 A类和B类的分类目的只是指出评定不确定度分量的两种不同方法,并且只是为了讨论的方便。这种分类不是指这两类分量本身性质上有差别。两类评定都基于概率分布,并且由两类评定得到的不确定度分量都是用方差或标准差定量表示。

B.1.3.5 表征用A类评定得到的不确定度分量的估计方差 u^2 是由一系列重复观测值计算得到的,它就是熟知的统计方差估计值 s^2 (见4.2)。标准差的估计值 u 是 u^2 的正平方根,因此 $u=s$,为方便起见,有时称为“A类标准不确定度”。由B类评定得到的不确定度分量,其方差估计值 u^2 是根据已有的信息评定的(见4.3),其标准差的估计值 u 有时称为“B类标准不确定度”。

因此,A类标准不确定度是由一组观测得到的频率分布导出的概率密度函数得到。而B类标准不确定度是由一个假定的概率密度函数得到,此函数基于对一个事件发生的信任程度[通常称为主观概率]。这两种方法都采用了公认的对概率的解释。

注:不确定度分量的B类评定通常基于一组相对可靠的信息(见4.3.1)。

B.1.3.6 当测量结果是由若干个其他量值得到时,测量结果的标准不确定度称为合成标准不确定度,用 u_c 表示。它是测量结果的标准差的估计值,并等于由所有方差和协方差分量求得的合成方差的正平方根,用本标准中称为不确定度传播律的方法进行评定(见第5章)。

B.1.3.7 为满足工业和商业应用,以及健康和安全领域的要求,将合成标准不确定度 u_c 乘以包含因子 k 得到扩展不确定度 U 。其目的是提供测量结果的一个区间,可期望被测量值分布的大部分包含在该区间内。因子 k 的选取要根据区间所需的包含概率或置信水平而定,一般在2~3范围内(见第6章)。

注:应声明包含因子 k ,以便能够重新得到被测量的标准不确定度,因为对于其他依赖于该量的测量,计算测量结果的合成标准不确定度时会用到。

B.1.4 实际的考虑

B.1.4.1 如果改变测量结果所依赖的各个量,就可以用统计方法评定其不确定度。然而由于时间和资源有限,实际做到的可能性很小。测量结果的不确定度通常用测量的数学模型和不确定度传播律来评定。本标准中,假设测量能被数学模型化,其模型化程度与测量所需的准确度要求相应。

B.1.4.2 由于数学模型可能不完善,所有相关量应该尽可能充分地按实际情况变化,以便不确定度评定能够基于尽可能多的观测数据之上。可行时,为可靠地评定不确定度,应该根据长期的定量数据建立测量的经验模型,并使用核查标准和控制图指明测量过程是否处于统计控制。当观测数据,包括对同一被测量多次独立的测量结果,证明模型不完善时,应该修改数学模型。良好设计的实验方案也非常有助于对不确定度进行可靠地评定,并且其本身是测量技巧的重要部分。

B.1.4.3 为判断测量系统是否正常运行,通常要把实验观测到的输出量的变异性,也就是实验标准差与表征该测量结果的各项不确定度分量合成得到的预期的标准差作比较。这种情况下,只需考虑对输出值的实验观测变异性有贡献的分量(不论由A类还是B类评定得到)。

注:把对变异性有贡献和没有贡献的分量分成两组并适当标记将有助于进行分析。

B.1.4.4 有些情况下,测量结果的不确定度评定中可不必包括为补偿系统影响而使用的修正值的不确定度。尽管已经评估了其不确定度,如果它对测量结果的合成标准不确定度没有显著贡献,在评定不确定度时可以忽略不计。如果修正值本身与合成标准不确定度相比不显著时,修正值也可忽略不计。

B.1.4.5 在实际工作中,尤其在法制计量领域中,经常发生这样的情况:对被检样品的测量是通过与一个测量标准比较而进行的,相对于测量所需的准确度来说,测量标准及比较程序(方法)引入的不确定度可以忽略。例如,用一组已校准的标准砝码去校准商用秤的准确度,在这种情况下,测量可被视为是要

确定被校仪器的误差,因为不确定度的分量非常小而可以忽略(见 D.2.4.2)。

B.1.4.6 由测量结果提供的被测量的估计值有时可依据所用的测量标准的值来表示,而不是按国际单位制的有关单位来表示。这种情况下,测量结果的不确定度比用国际单位制有关单位表示结果时明显地小(实际上被测量已经被重新定义为测量值与标准值之比)。

示例:

一台高质量的齐纳电压标准通过与约瑟夫森效应电压基准进行比对校准,该基准是基于 CIPM 向国际推荐采用的约瑟夫森常数的约定值。当按约定值报告时,齐纳电压标准的已校准电压 V_s 的相对合成标准不确定度 $u_c(V_s)/V_s$ (见 5.1)为 2×10^{-8} 。但当 V_s 按电压 V 的国际单位制单位报告时, $u_c(V_s)/V_s$ 为 4×10^{-7} ,因为其中包括了约瑟夫森常数转换为国际单位制的值的不确定度。

B.1.4.7 记录或分析数据时的失误会给测量结果引入明显的未知误差。较大的失误通常可以通过对数据的适当审查来识别,但小的失误容易被掩盖,甚至被当作随机变化。不确定度的评定不考虑这种失误。

B.1.4.8 虽然本标准提供了评估不确定度的框架,但不能替代周密的思考、理智的诚实和专业的技巧。不确定度的评定既不是例行的任务,也不是纯粹的数学,它取决于对测量和被测量的本质的充分理解和认识。所以测量结果的不确定度评定的质量和实用性主要取决于对赋值有贡献的影响量认识的程度、细致的分析和齐全性。

B.1.4.9 本标准中使用的术语“测量结果”是基于 VIM 第 2 版的定义“由测量所得到的赋予被测量的值”,测量结果主要是指测得的量值,虽然 VIM 第 2 版也说明在完整表述测量结果时应包含测量不确定度,但并没有规定测量不确定度是测量结果的一部分。在 VIM 第 3 版中,“测量结果”的定义为“与其他有用的相关信息一起赋予被测量的一组量值”,测量结果通常用单一测得的量值和对应的测量不确定度表示,强调了测量不确定度通常应作为测量结果的一部分,与测量结果同时给出。本标准没有根据 VIM 第 3 版的变化修改使用术语“测量结果”的内容,建议在使用本标准时注意。

B.2 “真”值、误差和测量不确定度

术语“真值”在传统的关于不确定度的出版物中一直在使用,但在本标准中不采用,其理由在本附录中阐述。

术语“被测量”,“误差”和“不确定度”经常被误解,因此本附录讨论了它们的基本概念。

本附录中展示了两张图(图 B.1、图 B.2),用图示法说明为什么在本标准中采用的不确定度概念是以测量结果以及评定的不确定度为基础的,而不是基于不可知量“真”值和误差。

B.2.1 被测量

B.2.1.1 被测量在 VIM 第 3 版中定义为“拟测量的量”,在 VIM 第 2 版和 IEC 60050-300:2001 中定义为“受到测量的特定量”。对被测量的说明要求了解量的类型,以及含有该量的现象、物体或物质状态的描述,包括任何有关成分及所涉及的化学实体。原则上说,没有无穷多信息时,被测量不可能被完整地描述。因此,被测量定义的不完整引入了测量结果的不确定度,该不确定度分量相对于测量所要求的准确度,可能很大也可能不大。

B.2.1.2 通常,对被测量的定义要规定其一定的物理状态和条件。

示例:

声音在由摩尔分数为 $N_2 = 0.780\ 8$, $O_2 = 0.209\ 5$, $Ar = 0.009\ 35$, $CO_2 = 0.000\ 35$ 成分组成的干燥空气中,温度 $T = 273.15\ K$ 和压力 $p = 101\ 325\ Pa$ 时的速度。

B.2.2 复现量

理想情况下,测量所复现的量应完全与定义的被测量一致。然而在通常情况下,定义的量是不可能复现的,测量是在被测量的近似量上进行的。

B.2.3 “真”值和已修正的值

B.2.3.1 即使复现量实际上完全满足被测量的定义,为了得到预期的测量结果,应修正复现量的测量结果与被测量之间的差异。对其他所有已别识的显著的系统影响也应进行修正。虽然最终的已修正结果有时被视为被测量的“真”值的最佳估计,但是实际上该结果只是被测量的量值的最佳估计。

B.2.3.2 举一个例子,假设被测量是某给定板材在规定温度时的厚度。将样品放至接近规定的温度,用千分尺测量其特定位置的厚度,在相同的测量位置、温度及千分尺所施加的压力下,测出的厚度就是复现量。

B.2.3.3 先要确定测量时的温度和所施加的压力,考虑千分尺的校准曲线、温度偏差及样品在千分尺夹持力下的微小变形的影响,对复现量进行修正。

B.2.3.4 已修正的测量结果可以称为“真”值的最佳估计值,这种意义上的“真”就可以认为是完全满足被测量定义的量值。但是千分尺在板材的不同部位测量时复现量就会不同,就会有不同的“真”值。然而,真值应与被测量的定义一致,因为后者不规定在板材的特定位置上测量厚度。因此在这种情况下,由于被测量的定义不完整,“真”值就有一个不确定度,可以通过在板材的不同位置上的测量来评定其不确定度。从某种意义上说,每个被测量都具有一个“本征”不确定度。原则上说这种不确定度是可以用某种方法估计的。这就是被测量可被测定的最小的不确定度。达到这样的不确定度的每个测量可以认为是被测量的最佳可能的测量。为了获得更小的不确定度,就要求对被测量作更加完整的定义。

注1:在本例中,被测量的说明还留有許多说明其他因素的余地。如大气压力、湿度、板材在重力场中的姿态、支撑方式等也可能对厚度有影响。

注2:虽然应当对被测量给出足够详细的定义,使由于定义不完整而引入的不确定度与要求的测量准确度相比可以忽略,但应认识到,这一点经常是不现实的。例如,由于假设某些影响可以忽略而没有规定其参数,实际上这种假设是不合理的;也可能隐含着某些不可能完全满足的条件及其实现的不完善有时很难被考虑到,由此导致定义可能是不完整的。举例:以 B.2.1.2 的例子而言,声音的速度意味着无穷多个幅度非常小的平面波,测量不满足这些条件的程度、衍射和非线性影响均需要考虑在内。

注3:对被测量的规定不充分会导致不同实验室的测量从表面看来是同样的量,其测量结果之间却不一致。

B.2.3.5 本标准中避免使用术语“被测量的真值”或“量的真值”(常缩写为“真值”),因为“真”字被视为多余的。“被测量”意味着“受测量的特定量”,因此“被测量的值”意思是“受测量的特定量的值”。由于“特定量”通常理解为有限的或规定的量,在“被测量的真值”(或“量的真值”)中的形容词“真”是不必要的——被测量(或量)的“真”值就是被测量(或量)的值。此外,正如前面讨论所指出的,“真值”这一词仅是一个理想的概念。

B.2.4 误差

一个已修正的测量结果不是被测量的值——也就是说因为复现量的测量不完善而存在着误差。观测值的随机变化(随机影响)、系统影响的修正值确定的不充分和对某个物理现象的认识不充分(也是系统影响)都会使测量不完善。无论是复现量的值还是被测量的值都不可能准确知道,所能知道的只是它们的估计值。在前面的例子中测得的板材的厚度可能存在误差,也就是说可能与被测量的值(板材的厚度)存在差异,因为以下因素均可能对测量结果的未知误差有贡献:

- a) 当千分尺重复用于测量同一复现量时,千分尺示值的微小差异;

- b) 千分尺校准的不完善；
- c) 温度和所施加压力测量的不完善；
- d) 对温度、大气压力和湿度对样品或千分尺的影响的认识不充分。

B.2.5 不确定度

B.2.5.1 尽管导致测量误差的贡献量的准确值未知或不可知,但与引起误差的随机影响和系统影响有关的不确定度是可以评定的。但是即使评定的不确定度很小,仍然不能保证测量结果的误差很小。可能由于认识不到而忽略某个系统影响,因此测量结果的不确定度并不一定能表明测量结果接近被测量值的程度,它只不过是现有可利用的知识下与最佳值接近程度的一种估计。

B.2.5.2 测量不确定度表明了一个事实,对给定的被测量和给定的该被测量的测量结果,存在的不是一个值,而是分散在测量结果附近的无穷多个值,这些值是与所有的观测值和数据以及人们对物理世界的认识相一致的,并按不同的置信度赋予被测量的。

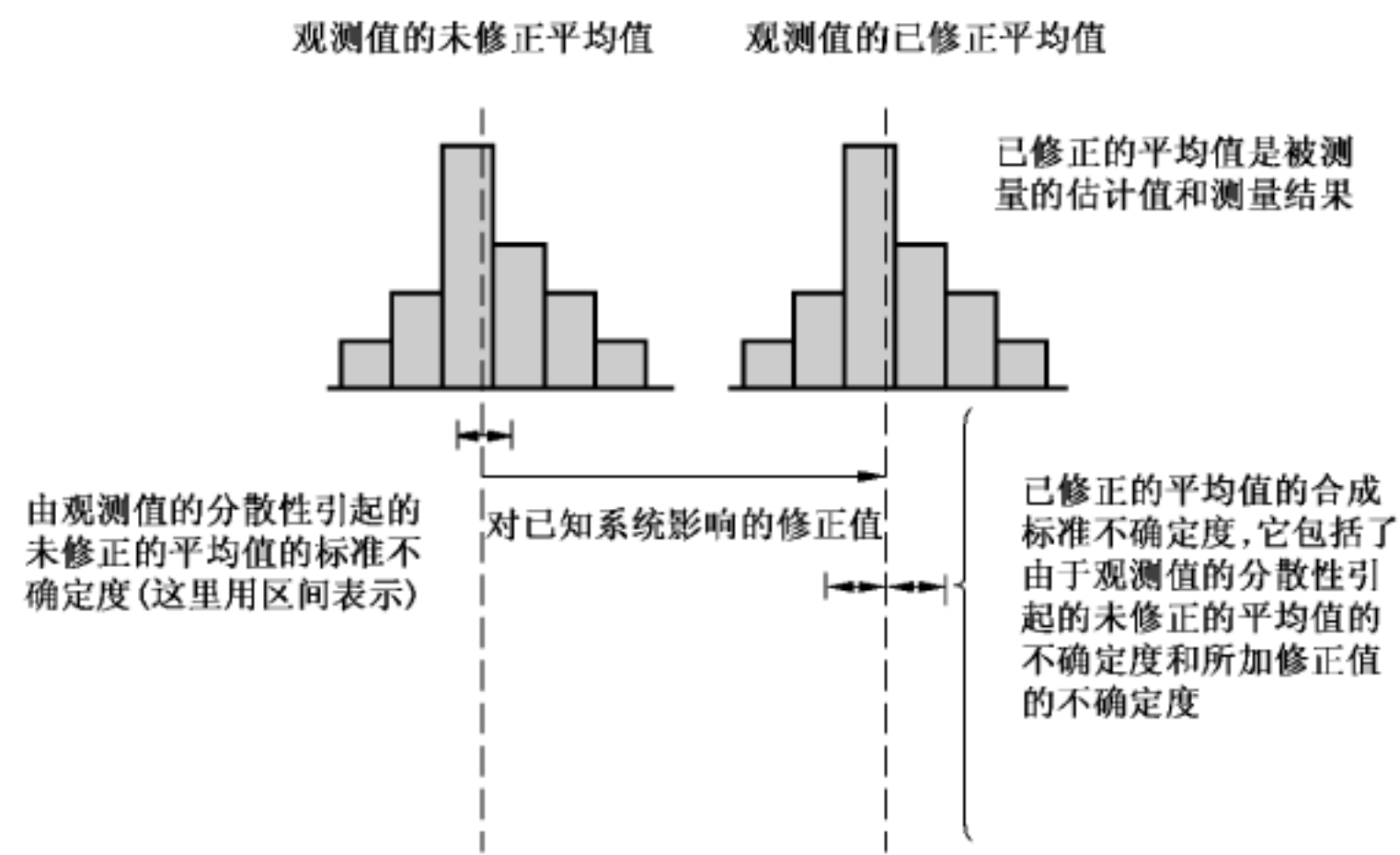
B.2.5.3 幸运的是在许多实际测量情况中,本附录讨论的许多情况无需考虑,例如当被测量被很好地定义时;当标准或仪器是用可溯源到国家标准的公认的参考标准校准时;以及当校准修正值的不确定度与仪器示值的随机影响或有限的观测次数引入的不确定度相比不显著时(见 C.4.3)。然而,对影响量及它们的影响认识不足常会对测量结果的不确定度有显著的贡献。

B.2.6 图解说明

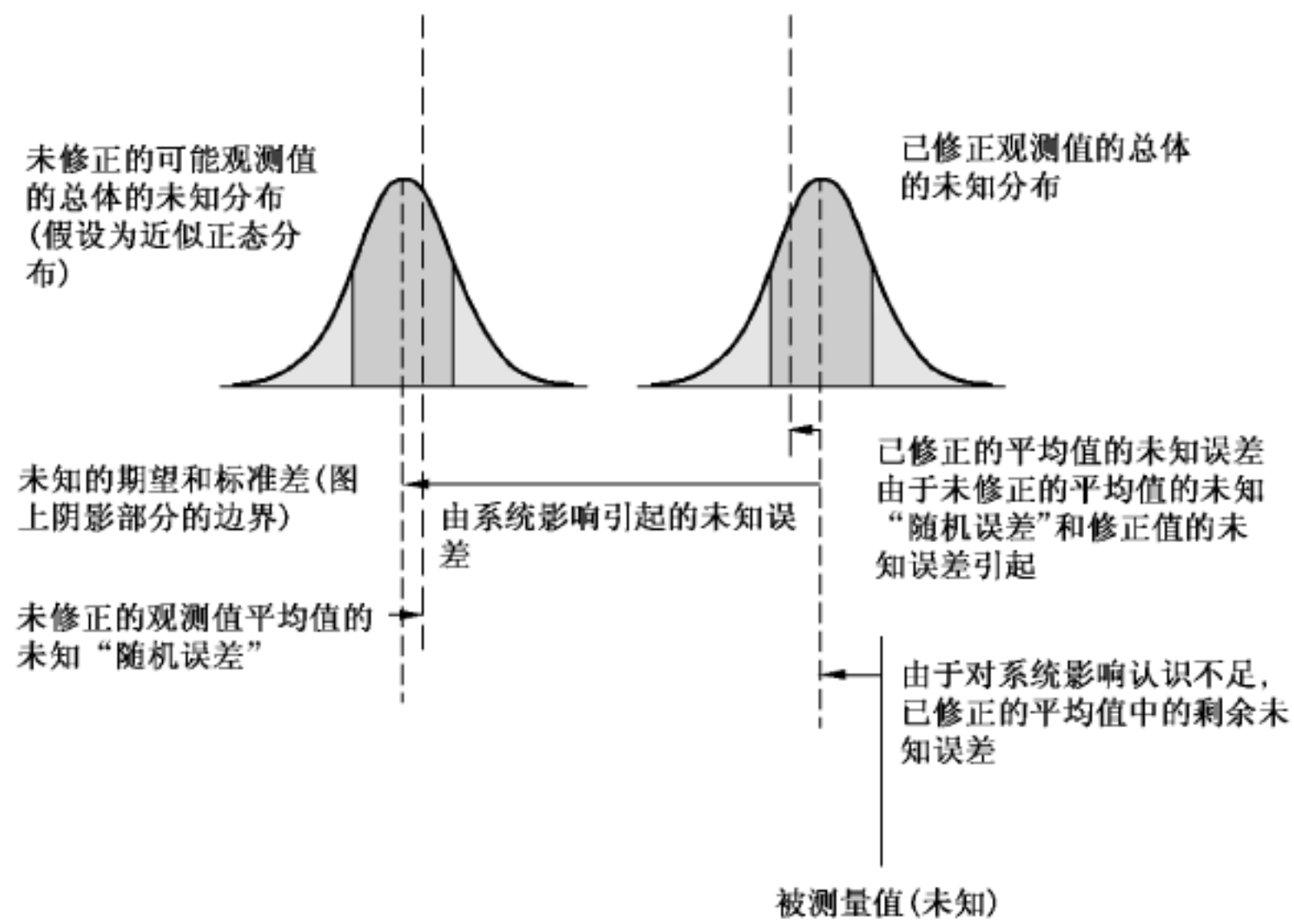
B.2.6.1 图 B.1 描述了本标准中的一些概念。它说明了为什么本标准把重点放在不确定度而不是误差上。测量结果的确切的误差通常是未知的和不可知的。人们能做的是估计输入量的值,包括对已认识的系统影响的修正值及其标准不确定度(估计标准差),既可以从重复观测的样本由未知的概率分布来估计,也可以根据可获得的信息假设的主观或先验的概率分布来估计,然后由各输入量的估计值计算出测量结果,并由这些估计值的标准不确定度计算出测量结果的合成标准不确定度。只有在确信这些工作做得很合理,没有忽略显著的系统影响时,才能认为测量结果是被测量的值的可靠估计,并且其合成标准不确定度是其可能误差的可靠度量。

注 1: 在图 B.1a)中为了说明问题,观测值用直方图表示[见 4.4.3 和图 1b)]。

注 2: 误差的修正值等于负的误差估计值。在图 B.1 和图 B.2 中说明误差修正值的箭头与说明误差的箭头的长度相同,但方向相反,反之亦然。图中文字清楚指出特定的箭头是指修正值还是误差。



a) 基于观测量的概念



b) 基于未知量的理想概念

图 B.1 值、误差和不确定度的图解说明

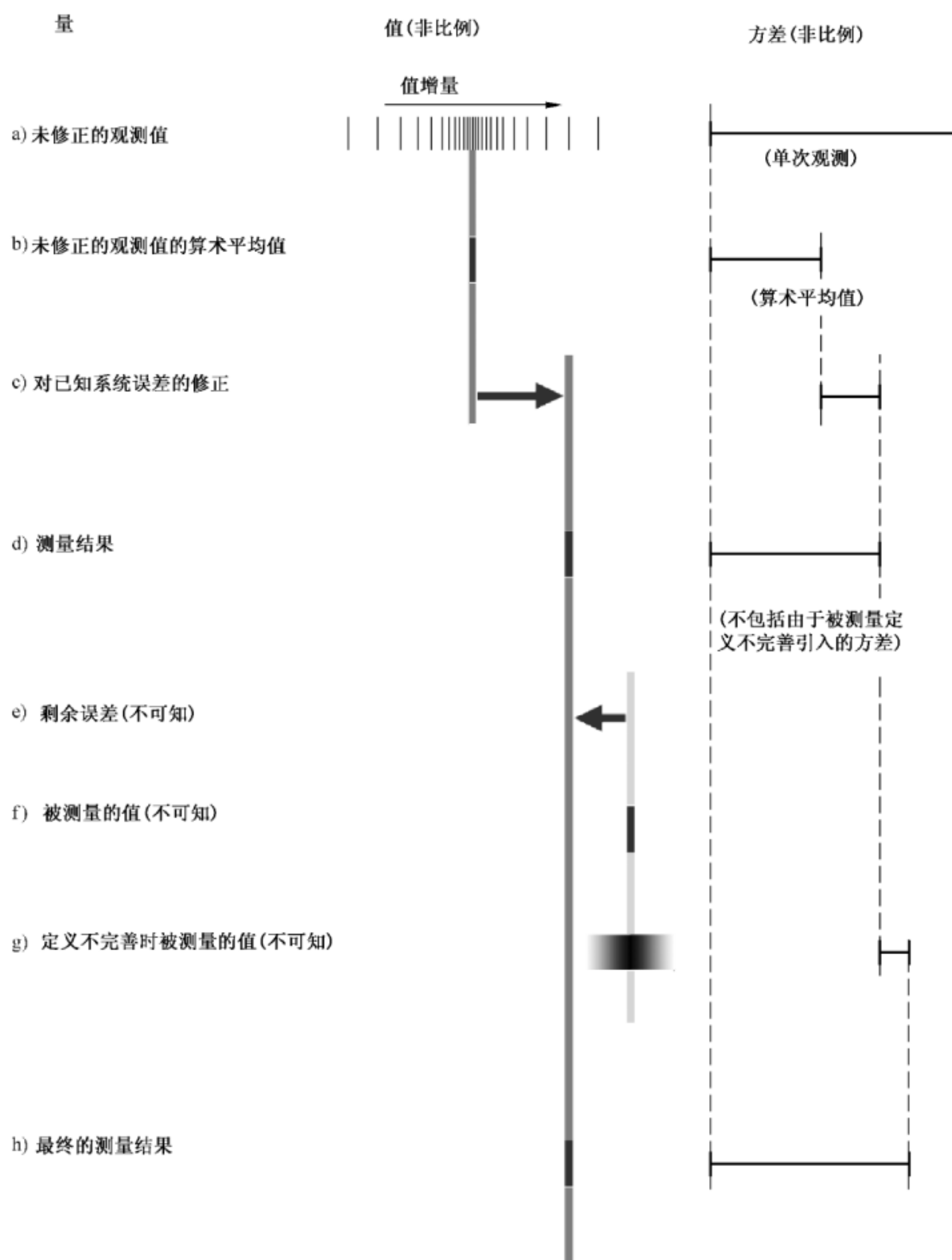


图 B.2 值、误差和不确定度的图解说明

B.2.6.2 图 B.2 使用不同的方式描述了图 B.1 说明的那些概念。该图还描述了一种概念：如果被测量的定义不完整就可能有许多被测量的值[图 B.2 中的 g)]。当用方差度量由定义不完整引起的不确定度时，同样用对被测量用同一方法、仪器等进行多次测量来评定(见 B.2.3.4)。

注：在以“方差”为标题的一栏中，方差可理解为 5.1.3 中式(11a)定义的方差 $u^2(y)$ ，因此如图所示它们是线性相加的。

B.3 统计学术语和概念的解释

B.3.1 期望 expectation

概率密度函数为 $p(z)$ 的随机变量 z 的期望 $g(z)$, 则:

$$E[g(z)] = \int g(z)p(z)dz$$

从概率密度函数 $p(z)$ 的定义可知, $\int p(z)dz = 1$ 。

用 μ_z 表示随机变量 z 的期望, 它也称为 z 的期望值或平均值, 则:

$$\mu_z = E(z) = \int zp(z)dz$$

z 的算术平均值或均值是:

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$$

B.3.2 方差 variance

随机变量的方差是其期望值的二次偏差的期望。因此, 概率密度函数为 $p(z)$ 的随机变量 z 的方差为:

$$\sigma^2(z) = \int (z - \mu_z)^2 p(z) dz$$

其中, μ_z 表示随机变量 z 的期望。方差 $\sigma^2(z)$ 可以用下式评估:

$$s^2(z_i) = \frac{1}{n-1} \sum_{j=1}^n (z_j - \bar{z})^2$$

其中, $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$, z_i 是 z 的第 n 次独立观测值。

注 1: $s^2(z_i)$ 的表达式中的因子 $n-1$ 来自 z_i 与 \bar{z} 的关联, 并反映集合 $\{z_i - \bar{z}\}$ 只有 $n-1$ 个独立项目的事实。

注 2: 如果随机变量 z 的期望 μ_z 已知, 则方差用下式评估:

$$s^2(z_i) = \frac{1}{n} \sum_{i=1}^n (z_i - \mu_z)^2$$

算术平均值或均值的方差, 而不是单一观测值的方差, 是测量结果的不确定度恰当表征, 应仔细区分是变量的方差还是其均值的方差。 z 的 n 个独立观测值 z_i 的算术平均值的方差为 $\sigma^2(\bar{z}) = \sigma^2(z_i)/n$, 实验标准差为:

$$s^2(\bar{z}) = \frac{s^2(z_i)}{n} = \frac{1}{n(n-1)} \sum_{i=1}^n (z_i - \bar{z})^2$$

B.3.3 标准差 standard deviation

标准差是方差的正平方根。一个 A 类标准不确定度是由统计方法评估的方差的平方根获得的, 而通常, 确定一个 B 类不确定度更方便, 首先用非统计方法评估一个等效的标准差, 然后通过对这个标准差求开方获得等效的方差。

B.3.4 协方差 covariance

两个随机变量的协方差是对它们之间相互依赖程度的估量。随机变量 y 和 z 的协方差定义为:

$$\text{cov}(y, z) = \text{cov}(z, y) = E\{[y - E(y)][z - E(z)]\}$$

导出：

$$\begin{aligned} \text{cov}(y, z) &= \text{cov}(z, y) \\ &= \iint (y - \mu_y)(z - \mu_z)p(y, z)dydz \\ &= \iint yz p(y, z)dydz - \mu_y\mu_z \end{aligned}$$

其中 $p(y, z)$ 是变量 y 和 z 的联合概率密度函数。协方差 $\text{cov}(y, z)$ [也可用 $v(y, z)$ 表示] 可由 y 和 z 的 n 对独立的并对应的 y_i 和 z_i 的观测值, 通过 $s(y_i, z_i)$ 计算：

$$s(y_i, z_i) = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})(z_i - \bar{z})$$

$$\text{其中, } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$$

注：均值 \bar{y} 和 \bar{z} 的协方差可以由 $s(\bar{y}, \bar{z}) = s(y_i, z_i)/n$ 计算。

B.3.5 协方差矩阵 covariance matrix

对一个多变量的概率分布, 这些变量的方差和协方差为元素的矩阵 V 称为协方差矩阵。对角元素 $v(z, z) \equiv \sigma^2(z)$ 或 $s(z_i, z_i) \equiv s^2(z_i)$ 为方差, 非对角元素 $v(y, z)$ 或 $s(y_i, z_i)$ 为协方差。

B.3.6 相关系数 correlation coefficient

相关系数是衡量两个变量的相互依赖关系的量, 它等于这两个变量的协方差与它们的方差的乘积的正平方根之比。因此, 由

$$\rho(y, z) = \rho(z, y) = \frac{v(y, z)}{\sqrt{v(y, y)v(z, z)}} = \frac{v(y, z)}{\sigma(y)\sigma(z)}$$

计算出：

$$r(y_i, z_i) = r(z_i, y_i) = \frac{s(y_i, z_i)}{\sqrt{s(y_i, y_i)s(z_i, z_i)}} = \frac{s(y_i, z_i)}{s(y_i)s(z_i)}$$

相关系数是一个数值为 $-1 \leq \rho \leq +1$ 或 $-1 \leq r(y_i, z_i) \leq +1$ 的纯数。

B.3.7 独立 independence

如果两个随机变量的联合概率分布是它们个体概率分布的乘积, 则这两个随机变量在统计学上是独立的。

注：如果两个随机变量是独立的, 则其协方差和相关系数为零, 但反之则不一定。

B.3.8 t 分布(学生分布) the t -distribution; student's distribution

t 分布或学生分布是连续随机变量 t 的概率分布, 其概率密度函数为

$$p(t, \nu) = \frac{1}{\pi\nu} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}, -\infty < t < +\infty$$

其中, Γ 是伽马函数, $\nu > 0$ 。 t 分布的期望为零, 其方差为 $\nu/(\nu-2)$, 其中 $\nu > 2$ 。当 $\nu \rightarrow \infty$, t 分布趋向于 $\mu=0$ 和 $\sigma=1$ 的正态分布。

如果随机变量 z 是期望为 μ_z 的正态分布, 则变量 $(\bar{z} - \mu_z)/s(\bar{z})$ 的概率分布是 t 分布, 其中 \bar{z} 是 z 的 n 个独立观测值 z_i 的算术平均值, $s(z_i)$ 是这 n 个观测值的经验标准差, $s(\bar{z}) = s(z_i)/\sqrt{n}$ 是均值 \bar{z} 的自由度为 $\nu = n - 1$ 的经验标准差。

附 录 C

(资料性附录)

建议书 INC-1(1980)的目的和基础

本附录对标准所依据的由不确定度声明工作组起草的建议书 INC-1(1980)的目的和统计基础做了简略的介绍,更多信息见参考文献[1,2,11,12]。

C.1 “保险”“随机”和“系统”

C.1.1 本标准提出了评定和表示测量不确定度的一种广泛应用的方法。它提供了一个现实的而不是“保险”的不确定度值,依据的概念就是由随机影响引起的不确定度分量与对系统影响修正不完善引入的不确定度分量之间没有内在的差别。所以这种方法与老方法相比在以下两个方面有共同点。

C.1.2 第一个观念是报告的不确定度应该是“保险”的和“保守”的,意思是绝不能因为评定过小而出错。事实上,因为测量结果的不确定度的评定有问题,通常是过于慎重而偏大。

C.1.3 第二个观念是引起不确定度的影响量通常认为不是“随机的”就是“系统的”两种不同的性质;每种性质的不确定度各自按自己的方法合成并分别报告(或当需要单个数时,按某种规定的方法进行合成)。事实上,合成不确定度的方法通常是按满足保险的要求来设计的。

C.2 现实的不确定度评定的理由

C.2.1 当报告被测量的值时,应给出该值的最佳估计及该估计值的不确定度最佳评定,如果不确定度出错,通常不可能确定按什么方向出错才“保险”。低估不确定度可能引起过分信赖报告值,有时带来尴尬或甚至引起严重后果。对过于谨慎地放大不确定度也可能产生不希望有的影响,可能导致测量设备的用户去购买比他们所需的仪器更贵的仪器,或者可能引起将价格很高的产品不必要地报废了或者拒绝校准实验室的服务。

C.2.2 我们不否定测量结果的使用者使用自己的倍乘因子去乘不确定度以便获得扩展不确定度,该扩展不确定度定义了一个具有规定包含概率的区间,以满足使用者自己的需要。也不否定在某种情况下,提供测量结果的单位能常规地用乘一个因子得到类似的扩展不确定度以满足测量结果的特殊用户的需要。然而,这类因子(通常要说明)需要应用到由现实的方法评定的不确定度才行,只有在不确定度用现实的方法确定后,由扩展不确定度定义的区间才具有所需的包含概率,并且可以容易地反过来进行计算。

C.2.3 从事测量的人员通常应在其分析中纳入其他人的测量结果,这些测量结果都有自己的不确定度。在评定他们自己的测量结果的不确定度时,需要由其他地方得来的每个测量结果的不确定度的最佳值,而不是“保险”值。此外,应有一种符合逻辑且简单的方法,用这种方法可以将这些输入的不确定度与他们自己观测的不确定度合成,给出自己结果的不确定度。建议书 INC-1(1980)提供了这样的一种方法。

C.3 用同一方法处理所有不确定度分量的理由

本节讨论的重点是一个简单的例子,由这个例子说明在对测量结果的不确定度的评定中,本标准如何用完全相同的方法来处理由随机影响引起入的不确定度分量和由系统影响引入的不确定度分量,因

此是对本标准采用的和 C.1.1 引证的观点进行的举例说明,即所有的不确定度分量本质相同,处理方法一样。讨论的起点是简化的标准差传递的数学表达式的推导,在本标准中称为不确定度传播律。

C.3.1 设输出量 $z = f(w_1, w_1, \dots, w_N)$ 取决于 N 个输入量 w_1, w_2, \dots, w_N , 其中每个 w_i 用一个适当的概率分布描述, w_i 的期望 $E(w_i) \equiv \mu_i$, 当 z 相对于 μ_z 的小偏差在 1 阶泰勒级数中用 w_i 相对于 μ_i 的小偏差表示时, f 的表达式为:

$$z - \mu_z = \sum_{i=1}^N \frac{\partial f}{\partial w_i} (w_i - \mu_i) \quad \dots\dots\dots (C.1)$$

其中所有的高次项假设都可忽略,且 $\mu_z = f(\mu_1, \mu_2, \dots, \mu_N)$ 。偏差 $z - \mu_z$ 的平方为

$$(z - \mu_z)^2 = \left(\sum_{i=1}^N \frac{\partial f}{\partial w_i} (w_i - \mu_i) \right)^2 \quad \dots\dots\dots (C.2a)$$

可以写成:

$$(z - \mu_z)^2 = \sum_{i=1}^N \left(\frac{\partial f}{\partial w_i} \right)^2 (w_i - \mu_i)^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial w_i} \frac{\partial f}{\partial w_j} (w_i - \mu_i)(w_j - \mu_j) \quad \dots (C.2b)$$

偏差的平方 $(z - \mu_z)^2$ 的期望是 z 的方差,即 $E[(z - \mu_z)^2] = \sigma_z^2$, 因此由式(C.2b)得:

$$\sigma_z^2 = \sum_{i=1}^N \left(\frac{\partial f}{\partial w_i} \right)^2 \sigma_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial w_i} \frac{\partial f}{\partial w_j} \sigma_i \sigma_j \rho_{ij} \quad \dots\dots\dots (C.3)$$

其中 $\sigma_i^2 = E[(w_i - \mu_i)^2]$ 是 w_i 的方差, $\rho_{ij} = v(w_i, w_j) / (\sigma_i^2 \sigma_j^2)^{1/2}$ 是 w_i 和 w_j 的相关系数, 其中 $v(w_i, w_j) = E[(w_i - \mu_i)(w_j - \mu_j)]$ 是 w_i 和 w_j 的协方差。

注 1: σ_z^2 和 σ_i^2 分别是 z 和 w_i 的概率分布的 2 阶中心矩。概率分布可以用其期望、方差和高阶中心矩完整地表征。

注 2: 5.2.2 中的式(13)[以及式(15)]用于计算合成标准不确定度,除了式(13)是用方差、标准差和相关系数的估计值表示外,它和式(C.3)是相同的。

C.3.2 在传统术语中,式(C.3)通常称为“通用的误差传播律”,其更适用于用 $\Delta z = \sum_{i=1}^N (\partial f / \partial w_i) \Delta w_i$ 表示方式,其中 Δz 是 w_i [见式(C.8)]的小变化 Δw_i 引起的 z 的变化。实际上按本标准进行时,将式(C.3)称为“不确定度传播律”是合理的,因为它指出了如何从输入量 w_i 的不确定度合成给出输出量 z 的不确定度,输入量 w_i 的不确定度取为等于 w_i 的概率分布的标准差,输出量 z 的不确定度取为等于 z 的概率分布的标准差。

C.3.3 式(C.3)也用于标准差倍数的传递,将每个标准差 σ_i 用其倍数 $k\sigma_i$ 代替,且对每个 σ_i 使用相同的 k 值,输出量 z 的标准差用 $k\sigma_z$ 代替,然而它不适用于包含区间的传递。如果每个 σ_i 用一个量 δ_i 代替, δ_i 定义了一个具有给定包含概率 p 的区间, z 的结果量 σ_z 不一定是具有同样包含概率 p 的区间,除非所有的 w_i 都是正态分布的。在式(C.3)中没有关于 w_i 的概率分布为正态性的假设。更为特殊的情况,如果在 5.1.2 的式(10)中的每个标准不确定度 $u(x_i)$ 都是从独立重复观测值评定,并对某个 p 值(如 $p = 95\%$)根据自由度查到相应的 t 因子,用 t 因子乘标准不确定度,由此得到的估计值 y 的不确定度不是相应的 p 值的区间(见 E.3 和 E.4)。

注:历史上,使用正态分布假设,从重复观测值导出的不确定度分量,与简单从上限值和下限值导出的不确定度分量是有区别的,其理由之一就是使用式(C.3)传递包含区间时要求具有正态性。

C.3.4 就以下例子而言, z 仅与输入量 w 有关, $z = f(w)$, 其中 w 是由 n 个 w_k 值的平均值估计的; n 个值是由对随机变量 q 的 n 次独立重复观测值 q_k 得到的,与 q_k 的关系为:

$$w_k = \alpha + \beta q_k \quad \dots\dots\dots (C.4)$$

其中, α 是恒定的“系统”偏差或对每个观测值的共同偏移; β 是共同的比例因子。尽管在观测过程中偏移和比例因子是固定的,还是假设用一个先验概率分布来表征, α 和 β 是这些概率分布的期望的最佳估计。

w 的最佳估计值是算术平均值或平均值 \bar{w} , 由下式得到:

$$\bar{w} = \frac{1}{n} \sum_{k=1}^n w_k = \frac{1}{n} \sum_{k=1}^n (\alpha + \beta q_k) \quad \dots\dots\dots (C.5)$$

于是用 $f(\bar{w}) = f(\alpha, \beta, q_1, q_2, \dots, q_n)$ 来估计 z 量, 并且其方差 $\sigma^2(z)$ 的估计值 $u^2(z)$ 由式(C.3)得到。如果为简单起见, 假设 $z = w$, 使 z 的最佳估计为 $z = f(\bar{w}) = \bar{w}$, 则估计值 $u^2(z)$ 可以迅速求得。由式(C.5)注意到:

$$\begin{aligned} \frac{\partial f}{\partial \alpha} &= 1 \\ \frac{\partial f}{\partial \beta} &= \frac{1}{n} \sum_{k=1}^n q_k = \bar{q} \end{aligned}$$

且

$$\frac{\partial f}{\partial q_k} = \frac{\beta}{n}$$

将 α 和 β 的估计方差分别用 $u^2(\alpha)$ 和 $u^2(\beta)$ 表示, 假设每个观测值是不相关的, 由式(C.3)求得:

$$u^2(z) = u^2(\alpha) + \bar{q}^2 u^2(\beta) + \beta^2 \frac{s^2(q_k)}{n} \quad \dots\dots\dots (C.6)$$

其中 $s^2(q_k)$ 是依据 4.2.2 中式(4)计算得到的观测值 q_k 实验方差; $s^2(q_k)/n = s^2(\bar{q})$ 是平均值 \bar{q} 的实验方差[4.2.3 中式(5)]。

C.3.5 在传统术语中, 式(C.6)右边第 3 项称为对估计方差 $u^2(z)$ 的“随机”贡献, 因为它通常随着观测次数 n 的增加而减小, 而前面两项称为是“系统”贡献, 因为它们与 n 无关。

更重要的是, 在用一些传统的对测量不确定度处理方法中, 式(C.6)是有问题的, 因为没有区分哪些是由系统影响引起的不确定度, 哪些是由随机影响引起的不确定度。尤其是, 切不可将由先验概率分布得到的方差与基于频率的分布得到的方差合成, 因为概率的概念被认为只适用于在实质相同条件下能重复大量次数的事件, 一个事件的概率 p ($0 \leq p \leq 1$) 表示该事件发生的相对频率。

与以频率为基础的概率观点不同, 一种同等有效的观点认为: 概率是一个事件发生的可信程度的度量^[13,14]。例如, 假设某一个人有可能赢得一小笔钱 D , 并且他是一个理性的打赌者。如果对以下两种打赌的选择无偏袒之意, 则对事件 A 的发生的可信程度为 $p = 0.5$:

- 1) 如果事件 A 发生则赢得数额为 D 的金钱, 如果不发生则什么也没有;
- 2) 如果事件 A 不发生则赢得数额为 D 的金钱, 如果发生则什么也没有。

本标准依据的建议书 INC-1(1980)就隐含地采用了这种概率的观点, 因本标准认为用式(C.6)计算测量结果的合成标准不确定度是合适的方法。

C.3.6 本标准采用基于可信程度的概率理论, 以标准差(标准不确定度)和不确定度传播律[式(C.3)]作为评定和表示测量不确定度的基础, 有三个方面的优点:

- a) 不确定度传播律可方便地将一个测量结果的合成标准不确定度纳入使用该结果的另一个结果的合成标准不确定度之中;
- b) 合成标准不确定度可以作为计算不确定度区间的基础, 在实际应用中很方便与该区间所要求的包含概率对应;
- c) 评估不确定度时不必分为“随机”分量或“系统”分量(或以任何其他方式分类), 因为所有的不确定度分量都以同样的方法处理。

其中 c) 的好处更多, 因为这种分类往往是引起混淆的根源; 不确定度分量既不是“随机”的也不是“系统”的, 它的特性是由相应量的使用而定。或更正规地说, 是由该量在描述测量的数学模型中呈现的意义来定的。因此当相应量用于不同方面时, “随机”分量可能会成为“系统”分量, 反之亦然。

C.3.7 根据上面 c) 给出的理由, 建议书 INC-1(1980) 没有将不确定度分量分为“随机”和“系统”两类, 实际上, 在计算测量结果的合成标准不确定度时, 没有必要对不确定度分量进行分类, 因此不需要任何分类方案。然而, 由于方便的标记有时可有助于观点的交流和讨论, 建议书 INC-1(1980) 给出了按两种不同的不确定度分量评定方法来划分的分类方案, 即不确定度“A类”和“B类”(见 0.7)。

将不确定度分量评定的方法进行分类避免了与分量本身分类有关的主要问题, 即分量的分类取决于相应量的如何使用。然而, 按评估方法进行分类而不按分量进行分类并不妨碍为测量的特定目的将用两种方法评定得到的每个分量归入专门的组, 例如对复杂测量系统的输出值, 比较其实验观测值的变异性和理论预测值的变异性。

C.4 用标准差作为不确定度的度量

C.4.1 式(C.3)要求: 不管输入量的估计值的不确定度是如何得到的, 它应评定为标准不确定度, 也就是估计标准差。如果评定为某种“保险”的值, 就不能用于式(C.3)中。尤其是, 如果“最大误差限”(公认的最佳估计值的最大可能偏差)被用于式(C.3)中, 将导致不确定度具有一种难以确定的含义, 无法用于后续计算其他量的不确定度(见 C.3.3)。

C.4.2 当输入量的标准不确定度不能用大量重复观测结果的分析来评定时, 应根据已有知识采用一种概率分布, 然而这种分布可能与实际有一定差异, 但它不是无效或不现实的, 它是现有知识的体现。

C.4.3 基于重复观测的评定不一定就比其他方法好, 就 $s(\bar{q})$ 而言, 它是正态分布随机变量 q 的 n 次独立观测值 q_k 的均值的实验标准差[见 4.2.3 中式(5)]。 $s(\bar{q})$ 是一个统计量, 它是 $\sigma(\bar{q})$ 的估计值, $\sigma(\bar{q})$ 是 \bar{q} 的概率分布的标准差, 即是无穷多次重复测量得到的 \bar{q} 值的分布的标准差。 $s(\bar{q})$ 的方差 $\sigma^2[s(\bar{q})]$ 近似为:

$$\sigma^2[s(\bar{q})] \approx \sigma^2(\bar{q}) / (2\nu) \dots\dots\dots (C.7)$$

其中 $\nu = n - 1$ 是 $s(\bar{q})$ 的自由度(见 E.3.3)。因此, $s(\bar{q})$ 的相对标准差 $\sigma[s(\bar{q})] / \sigma(\bar{q})$ 可以用作 $s(\bar{q})$ 的相对标准不确定度的度量, 它近似为 $[2(n-1)]^{-1/2}$, 这是 \bar{q} 的“不确定度的不确定度”, 完全由于有限次取样这个纯统计学的原因引起, 可能相当大。在 $n = 10$ 次观测时为 24%。有关的值在表 C.1 中给出, 该表的数值表明, 在实际的 n 次测量时, 统计的估计标准差的标准差不可忽略, 所以可得出结论是标准不确定度的 A 类评定不一定比 B 类评定更可靠, 而且在许多实际测量情况下, 观测次数是有限的, 由 B 类评定得到的分量可能比由 A 类评定得到的分量更可信。

C.4.4 有种争议认为, 与特定的测量方法的应用相关联的不确定度是表征随机变量的统计参数, 存在一种“真正的系统影响”的例子, 其不确定度应用不同的方法来处理。具有未知的固定偏移就是一个例子, 它在用某种方法测量时每次测量的偏移都是一样的, 它是由于方法的原理本身可能存在不完善或某个基本假设的不完善引起的。如果知道可能存在这种偏移, 并且确信偏移值还相当大, 则可以用一种概率分布来描述, 根据有关它可能存在并且较大等信息直接建立概率分布。因此, 将概率认为是对事件将要发生的可信程度的度量, 那这样一个系统影响的贡献也包含在测量结果的合成标准不确定度之中, 它是用先验概率分布的标准不确定度评定的, 并用与其他输入量的标准不确定度相同的方法处理。

示例:

一个特定的测量程序的技术规范要求输入量由特定的级数展开式计算得到, 展开式的高次项是不能准确知道的。由于不能处理这些项引起的系统影响, 导致一个未知的固定偏移, 这个偏移是不能通过重复测量程序的取样得到的。因此, 如果严格按照以频率为基础的概率解释, 与这个影响有关的不确定度不能评定, 不能包括在最终测量结果的不确定度之中。然而, 用可信程度来解释概率时就允许将表征这种影响的不确定度由一个先验分布来评定(由可获得的与不准确知道的项的有关知识导出), 并且与任何其他不确定度一样, 包括在测量结果的合成标准不确定度的计算中。

表 C.1 $\sigma[s(\bar{q})]/\sigma(\bar{q})$, 对正态分布的随机变量 q 进行 n 次独立观测的平均值 \bar{q} 的实验标准差的标准差, 相对于平均值的标准差之比值^{a, b}

观测次数 n	$\sigma[s(\bar{q})]/\sigma(\bar{q})$ %
2	76
3	52
4	42
5	36
10	24
20	16
30	13
50	10

^a 这些值是由 $\sigma[s(\bar{q})]/\sigma(\bar{q})$ 的准确表示式计算得到的, 不是由近似式 $[2(n-1)]^{-1/2}$ 得到的。

^b 表达式 $\sigma[s(\bar{q})]/\sigma(\bar{q})$ 中, 分母 $\sigma(\bar{q})$ 是期望 $E[S/\sqrt{n}]$, 分子 $\sigma[s(\bar{q})]$ 是方差 $V[S/\sqrt{n}]$ 的平方根, 其中 S 是随机变量, 等于服从平均值为 μ 、方差为 σ^2 的正态分布的 n 个随机变量 X_1, \dots, X_n 的标准差:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}, \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

S 的期望和方差: $E(S) = \sqrt{\frac{2}{n-1}} \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]} \sigma, V[S] = \sigma^2 - E[S]^2$

其中, $\Gamma(x)$ 是伽马函数, 当 n 为有限值时 $E[S] < \sigma$ 。

C.5 对不确定度两种观点的比较

C.5.1 本标准的重点是测量结果及其评定的不确定度, 而不是不可知量的“真”值和误差(见附录 B)。从使用的观点看, 测量结果就是赋予被测量的值, 测量结果的不确定度是能合理地赋予被测量的值的分散性的度量。本标准实际上要将经常混淆的不确定度与不可知量——“真”值和误差, 区分开来。

C.5.2 从“真”值和误差的角度来解释不确定度传播律式(C.3)的推导可以了解这种关系。在这种情况下 μ_i 被看作未知量, 为输入量 w_i 的唯一真值, 每个 w_i 与“真值”的关系假设为 $w_i = \mu_i + \epsilon_i$, 其中 ϵ_i 是 w_i 的误差。每个 ϵ_i 的概率分布的期望假设为零, $E(\epsilon_i) = 0$, 其方差为 $E(\epsilon_i^2) = \sigma_i^2$ 。式(C.1)变为:

$$\epsilon_z = \sum_{i=1}^N \frac{\partial f}{\partial w_i} \epsilon_i \quad \dots\dots\dots (C.8)$$

其中, $\epsilon_z = z - \mu_z$ 是 z 的误差; μ_z 是 z 的“真”值。假设取 ϵ_z 的平方的期望, 就得到与式(C.3)形式一样的公式, 但其中 $\sigma_z^2 = E(\epsilon_z^2)$ 是 ϵ_z 的方差; $\rho_{ij} = v(\epsilon_i, \epsilon_j) / (\sigma_i^2 \sigma_j^2)^{1/2}$ 是 ϵ_i 和 ϵ_j 的相关系数, 其中 $v(\epsilon_i, \epsilon_j) = E(\epsilon_i \epsilon_j)$ 是 ϵ_i 和 ϵ_j 的协方差。因此方差和协方差与输入量的误差有关, 而不是与输入量本身相关联。

注: 假定概率被视为事件将要发生的可信程度的度量, 意味着系统误差可以用与随机误差相同的方法处理, 且 ϵ_i 可以代表任意一种误差。

C.5.3 实际上, 观点的差异并不会导致测量结果和其不确定度在数值上的不同。

首先, 在两种情况下, 输入量 w_i 的可获得的最佳估计值用于获得由函数 f 得到的 z 的最佳估计值; 如果最佳估计值被认为是赋予被测量的最可能的值或是它们的“真”值的最佳估计, 那么在计算中没有什么差别。

其次,因为 $\epsilon_i = w_i - \mu_i$, 并且 μ_i 是唯一的和确定的值, 其没有不确定度, ϵ_i 和 w_i 的方差和标准差是相同的。这就意味着在两种情况下, 用作标准差 σ_i 的估计值以便得到测量结果的合成标准不确定度的标准不确定度是相同的, 并得到同样的不确定度数值。再者, 如果标准不确定度被视为输入量的概率分布的分散性的度量或者视为该量的误差的概率分布的分散性的度量, 在计算上没有什么差别。

注: 如果没有做 C.5.2 的注的假设, 就不能应用本节的讨论, 除非所有的输入量的估计值及这些估计值的不确定度是由重复观测值的统计分析得到, 也就是从 A 类评定得到。

C.5.4 尽管基于“真”值和误差的方法产生的数字结果与本标准的方法的结果一样(C.5.2 中注的假设情况下), 本标准的不确定度的概念消除了误差和不确定度的混淆(见 B.2)。事实上, 本标准中的使用方法的重点是放在量的观测值(或估计值)和该量值的变异性, 完全不必提起误差。

附录 D

(资料性附录)

评定不确定度分量的实用指南

本附录给出了评定测量不确定度分量的其他建议,主要从实际使用考虑,它是对第 4 章中已经给出的建议的补充。

D.1 由重复观测值评定的不确定度分量:标准不确定度的 A 类评定

D.1.1 随机性和重复观测值

D.1.1.1 由重复观测值确定的不确定度与其他评定方法相比通常更为客观,并具有统计学的严格性等特点。若理解为对观测值仅能运用统计公式进行评定,以及认为这种评定方法无需运用判断都是不正确的。

D.1.1.2 首先,要问:“测量过程中重复观测值完全独立重复到什么程度?”。如果所有的观测值来自同一样品,且取样是测量过程的一部分,因为被测量是一种材料的特性(而不是该材料给定样品的特性),那么观测值不具有独立的重复性;应把不同样本间可能存在的差异引起的方差分量与在单个样本上重复观测的观测方差相加。

如果调零是测量过程中的一个步骤,那么每次重复测量都需要调零,即使在重复观测期间漂移可以忽略,因为调零过程存在潜在的统计上可以确定的不确定度分量。

同样,如果读取一个气压表上的数值,原则上应该对每一次重复测量读一次数(最好把它扰动一下再让它回到平衡状态),因为即使大气压恒定,还可能存在着示值或读数的变化。

D.1.1.3 第二,要问:所有假定为随机性的影响量是否确实是随机的,它们分布的均值和方差是否恒定,或在重复观测的周期内影响量的值是否有漂移?如果有足够多的观测次数,分别计算观测周期的前半周期和后半周期测量结果的算术平均值及它们的实验标准差,将两个平均值进行对比,以判别它们之间的差别在统计学上是否显著,由此判断是否存在随时间变化的影响。

D.1.1.4 如果实验室“公共服务”的值(如电源的电压和频率,水的压力和温度,氮气的压力等)是影响量,通常情况下它们的变化存在着不可忽略的非随机成分。

D.1.1.5 如果由于“噪声”的影响,示值的最低位数字在观测期间不断变化,有时很难让人选出该数字的值。最好采用某种方法在任意瞬间将示值锁定,然后记录下来。

D.1.2 相关性

本节中大部分讨论内容也适用于标准不确定度的 B 类评定。

D.1.2.1 两个输入量 X_i 和 X_j 估计值的协方差在以下情况下可以取为零或者忽略不计:

- a) X_i 和 X_j 不相关(随机变量,不是假设为不变的物理量——见 4.1.1 注 1)。例如,它们是在不同的独立实验中在不同的时间重复测量的,或者因为它们代表了独立进行的不同评定的结果量;
- b) X_i 或 X_j 中任一个量可以作为常数处理;
- c) 没有足够的信息评定 X_i 和 X_j 估计值的协方差。

注 1: 另一方面,在某些情况下,如 5.2.2 注 1 的参考电阻的例子,输入量很明显是完全相关的,其估计值的标准不确定度是线性合成的。

注 2: 不同的实验也可能不是独立的,例如,在各个实验中使用同一仪器测量的情况(见 D.1.2.3)。

D.1.2.2 两个重复并且同时观测的输入量是否相关可以用 5.2.3 中的式(17)来判断。例如,如果一个没有进行温度补偿或温度补偿欠佳的振荡器的频率为一个输入量,环境温度也是一个输入量。如果它

们是同时被观测的,振荡器的频率和环境温度的协方差计算结果可能显露出显著的相关性。

D.1.2.3 实际上,输入量间通常是相关的,因为在估计它们的值时使用了相同的测量标准、测量仪器、参考数据或甚至具有显著不确定度的测量方法。为不失一般性,假设两个输入量 X_1 和 X_2 的估计值用 x_1 和 x_2 来表示,它们取决于一组不相关变量 Q_1, Q_2, \dots, Q_L 。由此 $X_1 = F(Q_1, Q_2, \dots, Q_L), X_2 = G(Q_1, Q_2, \dots, Q_L)$, 尽管其中某些变量仅在一个函数中出现,在另一个函数中不出现,如果 $u^2(q_l)$ 是 Q_l 的估计值 q_l 的估计方差,则 x_1 的估计方差可由 5.1.2 的式(10)得到:

$$u^2(x_1) = \sum_{l=1}^L \left(\frac{\partial F}{\partial q_l} \right)^2 u^2(q_l) \quad \dots\dots\dots (D.1)$$

$u^2(x_2)$ 有类似的表达式。 x_1 和 x_2 的估计协方差为:

$$u(x_1, x_2) = \sum_{l=1}^L \frac{\partial F}{\partial q_l} \frac{\partial G}{\partial q_l} u^2(q_l) \quad \dots\dots\dots (D.2)$$

因为对一个给定的 l ,只有 $\partial F/\partial q_l \neq 0$ 和 $\partial G/\partial q_l \neq 0$ 的那些项对总和有贡献,如果 F 和 G 中不含有共同的变量,则协方差为零。

两个估计值 x_1 和 x_2 的相关系数 $r(x_1, x_2)$ 由 $u(x_1, x_2)$ [式(D.2)] 和 5.2.2 中式(14)确定,其中, $u(x_1)$ 由式(D.1)计算, $u(x_2)$ 由类似的表达式计算得到 [见 F.2.3 中的式(F.9)]。两个输入估计量的估计协方差也可能既有统计分量 [见 5.2.3 中的式(17)], 也有本节讨论的因素引起的分量。

示例 1:

在同一测量中,标准电阻 R_s 既用于确定电流 I 又用于确定温度 t 。电流是用一个数字电压表测量标准电阻两端的电位差来确定的;温度是用一个电阻电桥和标准电阻,测量出已校的电阻性温度传感器的电阻 $R_t(t)$ 而确定,在 $15\text{ }^\circ\text{C} \leq t \leq 30\text{ }^\circ\text{C}$ 范围内,温度传感器的温度-电阻的关系为 $t = aR_t^2(t) - t_0$, 其中 a 和 t_0 是已知的常数。因此,电流是根据关系式 $I = V_s/R_s$ 确定的,温度是根据关系式 $t = a\beta^2(t)R_s^2 - t_0$ 确定的,其中 $\beta(t)$ 是电桥测得的比值 $R_t(t)/R_s$ 。

因为只有 R_s 在 I 和 t 两个关系式中是公共的, I 和 t 的协方差由式(D.2)得到:

$$u(I, t) = \frac{\partial I}{\partial R_s} \frac{\partial t}{\partial R_s} u^2(R_s) = \left(-\frac{V_s}{R_s^2} \right) [2a\beta^2(t)R_s] u^2(R_s) = \frac{2I(t+t_0)}{R_s^2} u^2(R_s)$$

为简单起见,在这个例子中,对输入量和它的估计值用相同的字母表示。

为了获取协方差的数值,需要将测得的 I 和 t 的数值,以及标准电阻的校准证书中给出的 R_s 和 $u(R_s)$ 的值代入此表达式中, $u(I, t)$ 的单位为 $\text{A} \cdot \text{ }^\circ\text{C}$, 相对方差 $[u(R_s)/R_s]^2$ 的量纲为 1 (后者也就是所谓的无量纲量)。

进一步设 P 是与输入量 I 和 t 有关的量, $P = C_0 I^2 / (T_0 + t)$, 其中 C_0 和 T_0 是已知的常数且不确定度可忽略 [$u^2(C_0) \approx 0, u^2(T_0) \approx 0$]。根据 5.2.2 中式(13), 由 I 和 t 的方差及其协方差计算得到 P 的方差:

$$\frac{u^2(P)}{P^2} = 4 \frac{u^2(I)}{I^2} - 4 \frac{u(I, t)}{I(T_0 + t)} + \frac{u^2(t)}{(T_0 + t)^2}$$

方差 $u^2(I)$ 和 $u^2(t)$ 是对 $I = V_s/R_s$ 和 $t = a\beta^2(t)R_s^2 - t_0$ 应用 5.1.2 中式(10)得到的, 结果为:

$$u^2(I)/I^2 = u^2(V_s)/V_s^2 + u^2(R_s)/R_s^2$$

$$u^2(t) = 4(t+t_0)^2 u^2(\beta)/\beta^2 + 4(t+t_0)^2 u^2(R_s)/R_s^2$$

为简单起见,假定常数 t_0 和 a 的不确定度可以忽略。因为 $u^2(V_s)$ 和 $u^2(\beta)$ 可以通过对电压表和电阻电桥分别重复读数而确定,这些表达式可以很容易进行估算。当然,在确定 $u^2(V_s)$ 和 $u^2(\beta)$ 时,也要考虑仪器本身固有的不确定度和所采用测量程序的不确定度。

示例 2:

在 5.2.2 注 1 的例子中,设每个电阻的校准值用 $R_i = \alpha_i R_s$ 表示,由重复观测得到的比值 α_i 的标准不确定度为 $u(\alpha_i)$ 。进一步,对每个电阻器设 $\alpha_i \approx 1$, 并设 $u(\alpha_i)$ 对每个校准值实际上相同,即 $u(\alpha_i) \approx u(\alpha)$ 。则由式(D.1)和式(D.2)得到 $u^2(R_i) = R_s^2 u^2(\alpha) + u^2(R_s)$ 及 $u(R_i, R_j) = u^2(R_s)$ 。这意味着由 5.2.2 中的式(14)得到任意两个电阻 ($i \neq j$) 的相关系数为:

$$r(R_i, R_j) \equiv r_{ij} = \left\{ 1 + \left[\frac{u(\alpha)}{u(R_s)/R_s} \right]^2 \right\}^{-1}$$

由于 $u(R_s)/R_s = 10^{-4}$, 如果 $u(\alpha) = 100 \times 10^{-6}$, $r_{ij} \approx 0.5$; 如果 $u(\alpha) = 10 \times 10^{-6}$, $r_{ij} \approx 0.990$; 如果 $u(\alpha) = 1 \times 10^{-6}$, $r_{ij} \approx 1.000$ 。因此当 $u(\alpha) \rightarrow 0$ 时, $r_{ij} \rightarrow 1$, 且 $u(R_i) \rightarrow u(R_s)$ 。

注：一般来说，在与校准值比较时，如在本例中，已校准项的估计值是相关的，相关程度取决于比对的不确定度与参考标准的不确定度之比。正如经常发生的实际情况，与参考标准的不确定度相比，比对的不确定度可以忽略不计，相关系数等于+1，并且每个被校准项的不确定度与其标准的不确定度相同。

D.1.2.4 如果被测量 Y 所取决的一组输入量 X_1, X_2, \dots, X_N [见 4.1 中式(1)] 中把原来对两个或多个 X_i 为共有的量作为附加的独立输入量 Q_i 包括在内了，就不再需要引入协方差 $u(x_i, x_j)$ (为了完全建立 Q_i 与被影响量 X_i 之间的关系，可能有必要进行附加的测量)。然而，在有些情况下，保留协方差比增加输入量的数目更为方便。对同时重复观测值的观测协方差可以应用同样的处理过程 [见 5.2.3 中式(17)]，但是恰当的附加输入量的识别通常是特定的和非物理的。

示例：

在 D.1.2.3 的例 1 中，将 I 和 t 用 R_s 表示，从而导出 P 的表达式为：

$$P = \frac{C_0 V_s^2}{R_s^2 [T_0 + a\beta^2(t)R_s^2 - t_0]}$$

将输入量 I 和 t 用 V_s, R_s 和 β 代替，就避免了 I 和 t 之间的相关性。因为这些量是不相关的， P 的方差可以根据 5.1.2 中式(10)得到。

D.2 由其他方法评定的不确定度分量：标准不确定度的 B 类评定

D.2.1 B 类评定的需求

如果一个测量实验室有充分的时间和资源，可以对不确定度每一个可想到的原因进行详尽的统计研究，例如，采用各种不同的仪器设备、不同的测量方法、测量方法的不同应用，以及测量理论模型的不同近似方式进行研究。与所有这些原因有关的不确定度就可以用一系列观测值的统计分析来评定，每一个原因可以用统计评定的标准差来表征。换言之，所有的不确定度分量都可由 A 类评定得到，但由于这种研究在实际中是一种不经济的方法，许多不确定度分量还是需要通过其他方法来评定。

D.2.2 数学方法确定分布

D.2.2.1 数字显示仪器的分辨力

数字仪器的不确定度来源之一是其显示装置的分辨力，例如，即使重复的示值都相同，贡献给重复性的测量不确定度仍然不为零，因为仪器的输入信号在一个已知的区间内变动，却给出同样的示值。如果显示装置的分辨力为 δ_x ，产生某一示值 X 的激励的值以等概率落在 $X - \delta_x/2$ 到 $X + \delta_x/2$ 区间内，该激励用宽度为 δ_x 的矩形概率分布来描述，方差为 $u^2 = (\delta_x)^2/12$ (参见 4.3.7 和 4.4.5)，意味着对任一示值，其标准不确定度为 $u = 0.29\delta_x$ 。

因此，对于一台称重仪器，其显示装置的最小有效显示数字为 1 g，由分辨力引入的方差为 $u^2 = (1 \text{ g})^2/12 = (1/12)\text{g}^2$ ，标准不确定度为 $u = (1/\sqrt{12})\text{g} = 0.29 \text{ g}$ 。

D.2.2.2 迟滞现象

某些类型的迟滞现象会引起类似的不确定度。迟滞是指测量仪器的指示值在连续上升或下降时可能与输入量之间存在一个固定和已知的差值。谨慎的操作者会记录下连续读数的方向，并采取适当的修正。但是迟滞现象的方向并不是总能被观测到：可能在仪器内在平衡点附近存在着振荡，使显示值取决于最终接近点的方向。如果由于这个原因引起的可能读数的范围为 δ_x ，则方差仍为 $u^2 = (\delta_x)^2/12$ ，由于迟滞引入的标准不确定度为 $u = 0.29\delta_x$ 。

D.2.2.3 有限精密度的计算

计算机对数据自动截断引起的数字舍入或截尾也是不确定度的一个来源。例如，一台字长为 16 位

的计算机,如果在计算过程中,一个具有这样字长的数字被另一个仅在第 16 位上有差别的数相减,只留下一位有效数字。这样的情况在“病态”算法的评估中会发生,并且是很难预测的。这类不确定度可以用实验确定,采用增加最重要输入量的微小增量,直到使输出量发生变化来计算(通常存在着一种与输出量大小成正比的输入量)。用这种方法获得的输出量的最小变化可以作为不确定度的度量,如果这个值是 δ_x ,那么方差为 $u^2 = (\delta_x)^2/12$, $u = 0.29\delta_x$ 。

注:可以用有限字长的计算机计算的结果与较大字长的计算机重复的同一计算结果相比较来验证不确定度的评定。

D.2.3 引入的输入值

D.2.3.1 一个输入量的引入值,不是从给定的测量过程中估计的,而是作为一种独立评定的结果以其他方式获得的值。通常这种引入值附有其不确定度的某种说明。例如,其不确定度可以是标准差、标准差的倍数或是具有所声明包含概率的一个区间的半宽度。也可以给出上限和下限,或者可能没有提供关于不确定度的信息。在后一种情况下,使用此值的人要用他们自己关于不确定度可能大小的知识,如:量的性质、来源的可靠性、这类量在实际中的不确定度等。

注:为方便起见,对引入的输入量的不确定度的讨论,放在本节中有关标准不确定度的 B 类评定中。这类量的不确定度可以由 A 类评定获得的分量组成,也可以由 A 类和 B 类两种评定获得的分量组成。由于计算合成标准不确定度并不需要区分两种不同方法评定的分量,故也不必知道一个引入量的不确定度的组成。

D.2.3.2 有些校准实验室采用上、下限来表示“不确定度”,此上下限定义了具有“最小”包含概率的一个区间,例如“至少”为 95%。这可以视为所谓“保险”的不确定度的一个例子(见 C.1.2),如果缺乏如何计算所需的知识,就不能转换得出标准不确定度。如果给了足够的信息,则可以根据本标准的规则进行重新计算;否则,不确定度应用可适用的方法进行独立评定。

D.2.3.3 有些不确定度用最大范围给出,量的所有值被认为都落在此范围内。通常在实际中假设在此范围内的所有值是等概率的(一种矩形概率分布),但如果理由期望接近边界的值比中心附近的值的可能性小时,不能假设成矩形分布。半宽度为 a 的矩形分布的方差是 $a^2/3$ 。包含概率为 99.73%、区间半宽度为 a 的正态分布的方差为 $a^2/9$ 。可以谨慎地采用两者的折中,如假设成三角分布,其方差为 $a^2/6$ (参见 4.3.9 和 4.4.6)。

D.2.4 测得的输入值

D.2.4.1 单次观测值,已校准的仪器

如果一个输入量的估计值是用一台特定的仪器单次观测得到的,并且这台仪器是由有很小的不确定度的标准校准过的,则估计值的不确定度主要是重复性。该仪器重复性测量的方差可以由以前的测量得到,以前测量的读数值不必与这次读数值是同一值,但应该足够接近,并可假设此方差可用于所研究的输入值。如果没有这样的信息可利用,则应根据测量装置或仪器的性质及其他类似结构仪器的已知方差做出估计。

D.2.4.2 单次观测值,已验证的仪器

并非所有的测量仪器都附有校准证书或校准曲线。然而,大多数仪器都按规定的标准建造,并由制造商或一个独立的机构验证,证明其符合该标准。通常该标准包含了计量要求,常以该仪器需要满足的“最大允许误差”的形式出现。用一个参考仪器进行比较的方法来确定该仪器是否符合要求,参考仪器的最大允许不确定度通常在标准中规定。于是此不确定度是被验证仪器的不确定度的一个分量。

如果被验证仪器的误差特性曲线不知道,就需要假设误差在允许范围内具有任意值,即具有矩形概率分布。然而,某些类型仪器的误差特性曲线在部分测量范围内是正的,而在另一部分是负的。有时这样的信息可以从标准和文献的研究中推断。

D.2.4.3 受控的量

测量经常在受控的标准条件下进行,在一组测量过程内假设条件恒定不变。例如,把样品放在一个搅拌过的油槽中进行测量,该油槽的温度是用一个恒温器控制的。在每次测量样品时可以测出油槽的温度,但是如果油槽的温度是循环的,那么样品的瞬时温度可能不是油槽温度计所指示的温度。根据热传导理论可以计算样品的温度波动及其方差,这超出了本标准范围,但计算应从油槽已知或假设的温度循环开始。用精密度较好的热电偶和一个温度记录器可以观测这种循环,但如果做不到这一点,也可以从控制特性推测出它的近似情况。

D.2.4.4 可能值的不对称分布

有时,量的所有可能值会落在单个极限值的一边。例如,当用压强计测量液柱的固定铅直高度 h (被测量)时,高度测量装置的轴可能偏离垂直方向一个小夹角 β 。由这个装置确定的距离 l 总是大于 h ,不可能比 h 小。这是因为 h 等于投影 $l \cos\beta$,即 $l = h / \cos\beta$;而 $\cos\beta$ 的值总是小于 1 的,不可能比 1 大。这种所谓的“余弦误差”也可能以同样的方式发生,使被测量 h' 的投影 $h' \cos\beta$ 等于观测到的距离 l ,即 $l = h' \cos\beta$,观测到的距离总是小于被测量的。

如果引入一个新的变量 $\delta = 1 - \cos\beta$,如同实际上常有的情况,假设 $\beta \approx 1$ 或 $\delta \ll 1$,存在两种不同的情况:

$$h = \bar{l}(1 - \delta) \dots\dots\dots (D.3a)$$

$$h' = \bar{l}(1 + \delta) \dots\dots\dots (D.3b)$$

这里 \bar{l} 是 l 的最佳估计值,它是 l 的 n 次独立重复观测值 l_k 的算术平均值,具有估计方差 $u^2(\bar{l})$ [见 4.2 中式(3)和式(5)]。因此,由式(D.3a)和(D.3b),为获得 h 和 h' 的估计值,就需要知道修正因子 δ 的估计值;而要获得 h 或 h' 估计值的合成标准不确定度,就需要知道 δ 的估计方差 $u^2(\delta)$ 。更准确地说,将 5.1.2 中的式(10)应用于(D.3a)和(D.3b)得到 $u_c^2(h)$ 和 $u_c^2(h')$ (符号分别为-和+)。

$$u_c^2 = (1 \mp \delta)^2 u^2(\bar{l}) + \bar{l}^2 u^2(\delta) \dots\dots\dots (D.4a)$$

$$\approx u^2(\bar{l}) + \bar{l}^2 u^2(\delta) \dots\dots\dots (D.4b)$$

为得到 δ 的期望值的估计值和 δ 的方差,假设在压强计中用于测量液柱高度的装置的轴固定在立体角内,倾斜角 β 值的分布是正态分布,其期望值为 0,方差为 σ^2 。尽管 β 可正可负,但是对所有的 β 值,其 $\delta = 1 - \cos\beta$ 是正的。如果不强制要求装置的轴对准,则轴的方向可在一个固定的角度内变化,由此也可能使方位角失准,此时 β 始终是正角。

在有约束或一维的情况下,概率元素 $p(\beta) d\beta$ 正比于 $\{\exp[-\beta^2/(2\sigma^2)]\} d\beta$;在不约束或二维的情况下,概率元素正比于 $\{\exp[-\beta^2/(2\sigma^2)]\} \sin\beta d\beta$ 分布。在两种情况下,都要用概率密度函数 $p(\delta)$ 的表达式去确定式(D.3)和式(D.4)用到的 δ 的期望值和方差。根据这些概率元素可以迅速得到它们,因为可以假设角 β 是很小的,因此 $\delta = 1 - \cos\beta$ 及 $\sin\beta$ 可以按 β 展开到最低阶,由此得到: $\delta = \beta^2/2$, $\sin\beta \approx \beta = \sqrt{2\delta}$, $d\beta = d\delta/\sqrt{2\delta}$ 。则概率密度函数为:

$$\text{一维时: } p(\delta) = \frac{1}{\sigma \sqrt{\pi\delta}} \exp(-\delta/\sigma^2) \dots\dots\dots (D.5a)$$

$$\text{二维时: } p(\delta) = \frac{1}{\sigma^2} \exp(-\delta/\sigma^2) \dots\dots\dots (D.5b)$$

其中,

$$\int_0^\infty p(\delta) d\delta = 1$$

两种情况下修正值 δ 的最可能值为零,根据式(D.5a)和式(D.5b)得到:一维情况下 δ 的期望和方差分别为 $E(\delta)=\sigma^2/2$ 和 $\text{var}(\delta)=\sigma^4/2$;二维情况下 δ 的期望和方差分别为 $E(\delta)=\sigma^2$ 和 $\text{var}(\delta)=\sigma^4$ 。式(D.3a)、式(D.3b)和式(D.4b)变为:

$$h = \bar{l}[1 - (d/2)u^2(\beta)] \quad \dots\dots\dots (D.6a)$$

$$h = \bar{l}[1 + (d/2)u^2(\beta)] \quad \dots\dots\dots (D.6b)$$

$$u_c^2(h) = u_c^2(h') = u^2(\bar{l}) + (d/2)l^{-2}u^4(\beta) \quad \dots\dots\dots (D.6c)$$

其中, d 是维数($d=1$ 或 2), $u(\beta)$ 是角度 β 的标准不确定度,取为假设正态分布的标准差 σ 的最佳估计值,并根据关于测量过程的所有可利用的信息来评定(B类评定)。这就是与输入量的不确定度有关的被测量估计值的一个例子。

虽然式(D.6a)到式(D.6c)是特定于正态分布,但这种分析也适用于 β 的其他分布。例如,假设 β 为对称矩形分布,在一维情况下其上下限为 $+\beta_0$ 和 $-\beta_0$,在二维情况下其上下限为 $+\beta_0$ 和 0 ;在一维情况下其期望和方差为 $E(\delta)=\beta_0^2/6$, $\text{var}(\delta)=\beta_0^4/45$;在二维情况下其期望和方差为 $E(\delta)=\beta_0^2/4$, $\text{var}(\delta)=\beta_0^4/48$ 。

注:这是函数 $Y=f(X_1, X_2, \dots, X_N)$ 按一阶泰勒级数展开的情况来得到 $u_c^2(y)$ 。因为 f 是非线性的, $\cos\beta \neq \cos\bar{\beta}$ (见 5.1.2 的注和 F.2.4),因此 5.1.2 中的式(10)不再适合。尽管分析可以完全以 β 来进行,但是引入变量 δ 可使问题简化。

另一个例子,反映了量的可能值全部落在单个极限值的一边的情形,是用滴定法测定溶液中成分的浓度,其终点由信号的触发来指示。所加试剂的量总是多于触发信号所需的量,从来不会少。超出极限点的超额滴定是数据分析中必需的变量。这种(和类似)情况下的程序是假设对超额量有一个适当的概率分布,并用它来确定超额量的期望值及方差。

示例:

假设超额量 z 为矩形分布,其下限为 0 ,上限为 C_0 ,则超额量的期望值是 $C_0/2$,方差为 $C_0^2/12$ 。如果超额量的概率密度函数取作 $0 \leq z < \infty$ 的正态分布,即 $p(z) = (\sigma \sqrt{\pi/2})^{-1} \exp[-z^2/(2\sigma^2)]$,期望为 $\sigma \sqrt{\pi/2}$ 方差为 $\sigma^2(1-2/\pi)$ 。

D.2.4.5 未修正校准曲线引起的不确定度

6.3.1 的注中讨论了这样一种情况,对于已知的显著系统影响的修正值 b 没有被加到测量结果中去,而是考虑到这种情况后加大了结果的“不确定度”。一个例子是用 $U+b$ 取代扩展不确定度 U ,其中 U 是假设 $b=0$ 时得到的扩展不确定度。有时在下述情况下应用这种方法:被测量 Y 是在参数 t 的一个范围内定义的,如温度传感器的校准曲线的情况; U 和 b 也取决于 t ;要在 t 的可能值范围内对被测量的所有估计值 $y(t)$ 给出“不确定度”单一值。在这些情况下,测量结果通常报告为: $Y(t) = y(t) \pm [U_{\max} + b_{\max}]$,其中下标“max”表示在所用的 t 值范围内最大的 U 值和最大的已知修正值 b 。

虽然本标准对已知的显著系统影响推荐将测量结果进行修正,但这不是总能行得通的,因为在计算和应用每个修正值,以及对每个 $y(t)$ 值计算和使用各自不确定度时,需要付出难以承受的代价。

解决该问题的一个比较简单的符合本标准原则的方法如下:

由下式计算修正值的平均值 \bar{b} :

$$\bar{b} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} b(t) dt \quad \dots\dots\dots (D.7a)$$

其中, t_1 和 t_2 是参数 t 的取值范围。取 $Y(t)$ 的最佳估计为 $y'(t) = y(t) + \bar{b}$,其中 $y(t)$ 是 $Y(t)$ 的最佳未修正估计值。在此取值范围内,平均修正值 \bar{b} 的方差为:

$$u^2(\bar{b}) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [b(t) - \bar{b}]^2 dt \quad \dots\dots\dots (D.7b)$$

上式没有考虑实际确定修正值 $b(t)$ 时的不确定度。由实际确定修正值 $b(t)$ 引起的平均方差为:

$$\overline{u^2[b(t)]} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} u^2[b(t)] dt \quad \dots\dots\dots (D.7c)$$

其中, $u^2[y(t)]$ 是修正值 $b(t)$ 的方差。同样, 由修正值 $b(t)$ 外的其他所有不确定度来源引起的 $y(t)$ 的平均方差为:

$$\overline{u^2[y(t)]} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} u^2[y(t)] dt \quad \dots\dots\dots (D.7d)$$

其中, $u^2[y(t)]$ 是由 $b(t)$ 外的其他所有不确定度来源引起的 $y(t)$ 的方差。于是, 对被测量 $Y(t)$ 的所有估计值 $y'(t) = y(t) + \bar{b}$ 的单个标准不确定度为下式的正平方根:

$$u_c^2[y'] = \overline{u^2[y(t)]} + \overline{u^2[b(t)]} + u^2(\bar{b}) \quad \dots\dots\dots (D.7e)$$

扩展不确定度 U 可以用 $u_c(y')$ 乘以一个适当选择的包含因子 k 得到, 即 $U = k u_c(y')$, 得到 $Y(t) = y'(t) \pm U = y(t) + \bar{b} \pm U$ 。然而, 对所有的 t 值使用同一个平均修正值而不是对每个 t 值进行修正时, 应清楚地说明 U 代表什么。

D.2.5 测量方法的不确定度

D.2.5.1 与测量方法有关的不确定度分量或许是最难评定的, 尤其是当测量方法对结果的变异性影响小于其他已知的因素时。但是, 可能还存在一些其他方法, 其中有些还是未知的或不切实际的, 它们对明显等效的结果会给出有系统性差异的结果。这就意味着与测量方法有关的不确定度一种先验概率分布, 而不是可以容易地取样并进行统计处理的分布。因此, 即使方法的不确定度占主要地位, 评价其标准不确定度所能获得的信息仅仅是人们现有的知识(见 C.4.4)。

注: 用不同的方法测定同一被测量, 既可在同一实验室也可在不同的实验室进行, 或在不同实验室内用相同的方法测量, 通常可以对某个特定方法对不确定度的贡献提供有价值的信息。总之, 实验室之间交换独立测量的测量标准或标准物质, 是评价不确定度结果的可靠性并识别以前未认识到的系统影响的很有效的方法。

D.2.6 样品的不确定度

D.2.6.1 许多测量是将被测样品与有类似特性的已知标准进行比较, 例如量块、砝码、电阻器、某些类型的温度计和高纯度材料等。在大多数情况下, 测量方法对样品的选择、样品的处理或各种环境影响量的影响不是特别敏感或有负面影响, 因为通常这些影响因素对被测样品和标准的影响(通常可以预测)是相同的。

D.2.6.2 一些实际测量中, 取样和样品处理对测量的影响很大。在对天然物质的化学分析中就常常如此。不像人工制造的材料具有可被证明的已超过测量所需的均匀性水平, 天然材料通常是很不均匀的。这种不均匀性导致两个附加的不确定度分量。第一个评定是要确定所选的样品代表被分析母体材料的充分程度, 第二个评定是要确定次要组分(未分析分量)影响测量的程度和测量方法处理这些次要组分的充分程度。

D.2.6.3 在有些情况下, 仔细设计实验后就有可能用统计方法来评定样品的不确定度(见 F.5 和 F.5.3.2)。然而, 通常当环境影响量对样品的影响较大时, 需要评定不确定度的人员具有必要的经验、技能和知识, 并能利用所有现有的信息。

附 录 E
(资料性附录)
自由度和置信水平(包含概率)

E.1 引言

E.1.1 本附录将讨论由被测量 Y 的估计值 y 与该估计值的合成标准不确定度 $u_c(y)$ 得到的扩展不确定度的问题,扩展不确定度 $U_p = k_p u_c(y)$,它定义了一个区间 $y - U_p \leq Y \leq y + U_p$,该区间具有一个很高的已规定的包含概率或置信水平 p 。这就涉及如何确定包含因子 k_p ,由 k_p 确定了测量结果 y 的一个区间,可期望它以规定的分数 p 包含了合理赋予被测量 Y 的值的分布的大部分(见第 6 章)。

E.1.2 在大多数实际测量情况下,具有规定包含概率的区间计算——实际上是在这种情况下对大多数的单个不确定度分量进行评定——仅仅是最佳的近似值。即使对用正态分布所描述的量进行多达 30 次重复观测,其平均值的实验标准差本身具有的不确定度约为 13%(见附录 C 中的表 C.1)。

在大多数情况下,试图去区别诸如后面示例中的区间是没有意义的,例如:一个包含概率为 95% 的区间(被测量 Y 的值 20 次测量有 1 次落在区间外)和一个 94% 或 96% 的区间(分别为 17 次中有 1 次或 25 次中有 1 次落在区间外)。要得到包含概率为 99%(100 次中有 1 次)和更高包含概率的合理区间是很难的,即使假设系统影响未被忽略,因为对输入量的概率分布的最极端部分或“尾巴”部分通常几乎没有信息。

E.1.3 为了得到对应于规定包含概率 p 的区间的包含因子 k_p 值,需要对测量结果及其合成标准不确定度所表征的概率分布有详细的了解。例如:用正态分布描述的一个量 z ,其期望值为 μ_z 、标准差为 σ ,则由 k_p 值得到一个区间 $\mu_z \pm k_p \sigma$,它包含了分布的分数 p ,因此具有包含概率或置信水平为 p ,此时的值可以迅速计算得到。表 E.1 给出了一些例子。

表 E.1 假设正态分布时具有包含概率 p 的区间的包含因子 k_p 值

包含概率 $p/\%$	包含因子 k_p
68.27	1
90	1.645
95	1.960
95.45	2
99	2.576
99.73	3

注: 对比一下,如果 z 是用一个期望值为 μ_z 、标准差 $\sigma = a/\sqrt{3}$ 的矩形概率分布描述, a 为分布的半宽度,包含概率 p 为 57.74% 时 $k_p = 1$;包含概率 p 为 $p = 95\%$ 时 $k_p = 1.65$;包含概率 p 为 99% 时 $k_p = 1.71$;包含概率 p 为 100% 时 $k_p \geq \sqrt{3} \approx 1.73$ 。从有限范围和无“尾巴”这个意义上来说,矩形分布比正态分布“窄”。

E.1.4 如果与被测量 Y 有关的输入量 X_1, X_2, \dots, X_N 的概率分布是已知的[它们的期望、方差和如果不是正态分布时的高阶矩(见 C.2.13 和 C.2.22),以及如果 Y 是输入量的线性函数: $Y = c_1 X_1 + c_2 X_2 + \dots + c_N X_N$,则 Y 值的概率分布可以由各个概率分布的卷积得到^[10]。于是相应于规定的包含概率 p 的区间的 k_p 值可以根据卷积结果的分布计算得到。

E.1.5 如果 Y 和它的输入量之间的函数关系是非线性的,用一阶泰勒级数展开近似是不够的(见 5.1.2

和 5.1.5), 则 Y 的概率分布不能由各输入量的分布的卷积得到。在这种情况下, 需要用其他分析方法或数学方法。

E.1.6 在实践中, 因为通常表征输入量概率分布的参数是估计值, 要准确知道一个给定区间的包含概率是不现实的, 并且因为对概率分布进行卷积的复杂性, 当需要计算具有规定包含概率的区间时, 很少进行概率分布的卷积运算。通常是利用中心极限定理进行近似。

E.2 中心极限定理

E.2.1 如果 $Y = c_1 X_1 + c_2 X_2 + \dots + c_N X_N = \sum_{i=1}^N c_i X_i$ 以及所有的 X_i 都是用正态分布表征的, 那么 Y 的卷积分布结果也是正态分布的。然而, 即使 X_i 值的分布不是正态的, 根据“中心极限定理”, Y 的分布通常可以用正态分布近似。这个定理还说明: 如果 X_i 值是独立的, 且 $\sigma^2(y)$ 值比任何非正态分布的 X_i 的单个分量 $c_i^2 \sigma^2(X_i)$ 都大得多, Y 的分布将是“近似正态”, 其期望值 $E(Y) = \sum_{i=1}^N c_i E(X_i)$, 方差 $\sigma^2(Y) = \sum_{i=1}^N c_i^2 \sigma^2(X_i)$, 其中 $E(X_i)$ 是 X_i 的期望, $\sigma^2(X_i)$ 是 X_i 的方差。

E.2.2 中心极限定理是重要的, 因为它表明了输入量的概率分布的方差在确定 Y 的卷积分布的形状时比分布的高阶矩起到了更重要的作用。此外, 它意味着当对 $\sigma^2(Y)$ 有贡献的输入量的数量增加时, 卷积分布收敛于正态分布。 $c_i^2 \sigma^2(X_i)$ 相互越接近 (相当于在实践中每个输入量的估计值 x_i 对被测量 Y 的估计值 y 的不确定度的贡献是可比拟的), 收敛就越快; X_i 的分布越接近正态, 则对 Y 来说成为正态分布所需的 X_i 的个数就越少。

示例:

矩形分布 (见 4.3.7 和 4.4.5) 是一个非正态分布的极端例子, 但即使只有三个这样等宽度的矩形分布, 其卷积也接近正态分布。如果每个矩形分布的半宽度是 a , 则每个的方差是 $a^2/3$, 卷积分布的方差 $\sigma^2 = a^2$ 。卷积分布的 95% 和 99% 包含概率的区间分别为 1.937σ 和 2.379σ , 而对标准差同为 σ 的正态分布的相应区间, 分别为 1.960σ 和 2.576σ (见表 E.1)^[10]。

注 1: 对包含概率 p 大于 91.7% 的每个区间, 正态分布的 k_p 值大于任意多个任意大小的矩形分布的卷积结果分布的相应值。

注 2: 根据中心极限定理, 具有期望 μ_q 和标准差 σ 的随机变量 q 的 n 次观测值 q_k 的算术平均值的概率分布, 不管 q 的概率分布是什么, 当 $n \rightarrow \infty$ 时, 趋近于期望值为 μ_q 和标准差为 σ/\sqrt{n} 的正态分布。

E.2.3 中心极限定理实践结果是: 当确定它的条件能近似满足时, 特别是如果合成标准不确定度 $u_c(y)$ 主要不是仅由少数观测值的 A 类评定得到的一个标准不确定度分量构成, 也主要不是由假设为矩形分布的 B 类评定得到的一个标准不确定度分量构成时, 使用合理的一阶近似, 计算扩展不确定度 $U_p = k_p u_c(y)$, 它提供了一个包含概率为 p 的区间, k_p 值可由正态分布得到。最常用的值在表 E.1 中给出。

E.3 t 分布和自由度

E.3.1 为了得到比 E.2.3 中所说的简单地由正态分布得到 k_p 值更好的近似, 应认识到计算一个具有规定的包含概率的区间需要的不是变量 $[Y - E(Y)]/\sigma(Y)$ 的分布, 而是变量 $(y - Y)/u_c(y)$ 的分布。这是因为在实践中经常得到的是 y , 即从 $y = \sum_{i=1}^N c_i x_i$ 得到的 Y 的估计值, 其中 x_i 是 X_i 的估计值; 并且 y 的

合成方差 $u_c^2(y)$ 由 $u_c^2(y) = \sum_{i=1}^N c_i^2 u^2(x_i)$ 评估, 其中 $u(x_i)$ 是估计值 x_i 的标准不确定度 (估计标准差)。

注: 严格来说, 在表达式 $(y - Y)/u_c(y)$ 中, Y 应该为 $E(Y)$ 。为简单起见, 本标准中仅在少数几个地方做这种区分。

总之, 对物理量、代表该量的随机变量和该变量的期望都用同一符号表示 (见 4.1.1 的注)。

E.3.2 如果 z 是一个正态分布的随机变量,期望为 μ_z ,标准差为 σ , \bar{z} 是 z 的 n 次独立观测值 z_k 的算术平均值, \bar{z} 的实验标准差为 $s(\bar{z})$ [见 4.2 中的式(3)和式(5)],则变量 $t = (\bar{z} - \mu_z) / s(\bar{z})$ 的分布称自由度为 $\nu = n - 1$ 的 t 分布或学生分布。

因此,如果被测量 Y 简单地为单个正态分布的量 $X, Y = X$;如果 X 是由 X 的 n 次独立重复观测值 X_k 的算术平均值 \bar{X} 估计的,其实验标准差为 $s(\bar{X})$,则 Y 的最佳估计值是 $y = \bar{X}$,估计值的实验标准差是 $u_c(y) = s(\bar{X})$ 。则:

$$t = (\bar{z} - \mu_z) / s(\bar{z}) = (\bar{X} - X) / s(\bar{X}) = (y - Y) / u_c(y)$$

是按照 t 分布而分布的。

$$P_r[-t_p(\nu) \leq t \leq t_p(\nu)] = p \quad \dots\dots\dots(\text{E.1a})$$

或者

$$P_r[-t_p(\nu) \leq (y - Y) / u_c(y) \leq t_p(\nu)] = p \quad \dots\dots\dots(\text{E.1b})$$

也可以写成

$$P_r[y - t_p(\nu)u_c(y) \leq Y \leq y + t_p(\nu)u_c(y)] = p \quad \dots\dots\dots(\text{E.1c})$$

在这些表达式中, $P_r[]$ 的意思是“概率”, t 因子 $t_p(\nu)$ 是给定自由度参数 ν 时的 t 值(见 E.3.3)——由 $-t_p(\nu)$ 到 $+t_p(\nu)$ 的区间包含了 t 分布的分数 p ,因此扩展不确定度:

$$U_p = k_p u_c(y) = t_p(\nu) u_c(y) \quad \dots\dots\dots(\text{E.1d})$$

确定了一个区间: $y - U_p$ 到 $y + U_p$,可方便地写为 $Y = y \pm U_p$,该区间可以预期包含了合理的表征 Y 值分布的分数 p ,其中 p 是该区间的包含概率或置信水平。

E.3.3 用 n 次独立观测值的算术平均值来估计单个量时,自由度 ν 为 $n - 1$,如 E.3.2 中所述。如果 n 次独立观测值用于确定最小二乘法拟合的直线的斜率和截距两个量,则它们各自的标准不确定度的自由度都为 $\nu = n - 2$ 。对由 n 个数据点拟合 m 个参数值的最小二乘法,其每个参数的标准不确定度的自由度是 $\nu = n - m$ (对自由度的进一步讨论见参考文献[15])。

E.3.4 根据不同的 ν 值和不同的 p 值选择的 $t_p(\nu)$ 值在表 E.2 中给出。当 $\nu \rightarrow \infty$ 时, t 分布趋于正态分布, $t_p(\nu) \approx (1 + 2/\nu)^{1/2} k_p$,在这个表达式中, k_p 是对一个正态分布变量获得具有包含概率为 p 的区间所需的包含因子。因此对一个给定的 p 在表 E.2 中的 $t_p(\infty)$ 等于表 E.1 中同样 p 时的 k_p 值。

注:通常 t 分布以分位数列表给出 $t_{1-\alpha}$ 的值,其中, $1 - \alpha$ 表示累积概率,定义如下:

$$1 - \alpha = \int_{-\infty}^{t_{1-\alpha}} f(t, \nu) dt$$

其中, f 是 t 分布的概率密度函数。因此 t_p 和 $t_{1-\alpha}(\nu)$ 由 $p = 1 - 2\alpha$ 相联系的。例如, $t_{0.975}$ 相应于 $1 - \alpha = 0.975$ ($\alpha = 0.025$) 时的 $t_{1-\alpha}(\nu)$ 值,等于 $p = 0.95$ 时的 $t_p(\nu)$ 值。

E.4 有效自由度

E.4.1 一般情况下,如果 $u_c^2(y)$ 是两个或更多个估计方差分量的合成: $u_i^2(y) = c_i^2 u^2(x_i)$ (见 5.1.3),即使每个 x_i 是正态分布输入量 X_i 的估计值,变量 $(y - Y) / u_c(y)$ 的分布也不是 t 分布。然而这个变量的分布可以用 t 分布近似,其有效自由度 ν_{eff} 由韦尔奇-萨特思韦特 (Welch-Satterthwaite) 公式得到^{[16][17][18]}:

$$\frac{u_c^4(y)}{\nu_{\text{eff}}} = \sum_{i=1}^N \frac{u_i^4(y)}{\nu_i} \quad \dots\dots\dots(\text{E.2a})$$

或者

$$\nu_{\text{eff}} = \frac{u_c^4(y)}{\sum_{i=1}^N \frac{u_i^4(y)}{\nu_i}} \quad \dots\dots\dots(\text{E.2b})$$

且

$$\nu_{\text{eff}} \leq \sum_{i=1}^N \nu_i \quad \dots\dots\dots (E.2c)$$

其中, $u_c^2(y) = \sum_{i=1}^N u_i^2(y)$ (见 5.1.3)。扩展不确定度 $U_p = k_p u_c(y) = t_p(\nu_{\text{eff}}) u_c(y)$, 由此提供了一个包含概率近似为 p 的区间 $Y = y \pm U_p$ 。

注 1: 如果由式(E.2b)得到的 ν_{eff} 值不是一个整数, 实际情况常常如此, 可用内插法或将 ν_{eff} 舍位到最近的一个较小的整数, 由表 E.2 查到相应的 t_p 值。

注 2: 如果一个输入量的估计值 x_i 本身是由两个或更多个其他估计值得到的, 则在式(E.2b)的分母中与 $u_i^4(y) = [c_i^2 u^2(x_i)]^2$ 一起使用的 ν_i 的值是由相当于式(E.2b)的表达式计算得到的有效自由度。

注 3: 根据测量结果的潜在用户的需要, 除 ν_{eff} 外, 分别处理由 A 类和 B 类评定得到的标准不确定度时可按式(E.2b)计算和报告 ν_{effA} 和 ν_{effB} 值。如果要单独表明 A 类和 B 类评定得到的标准不确定度对 $u_c^2(y)$ 的贡献, 可分别用 $u_{cA}^2(y)$ 和 $u_{cB}^2(y)$, 这些量的关系为:

$$u_c^2(y) = u_{cA}^2(y) + u_{cB}^2(y)$$

$$\frac{u_c^4(y)}{\nu_{\text{eff}}} = \frac{u_{cA}^4(y)}{\nu_{\text{effA}}} + \frac{u_{cB}^4(y)}{\nu_{\text{effB}}}$$

示例:

设 $Y = f(X_1, X_2, X_3) = bX_1 X_2 X_3$, 及正态分布输入量 X_1, X_2, X_3 的估计值 x_1, x_2, x_3 分别为 $n_1 = 10, n_2 = 5, n_3 = 15$ 次独立重复观测值的算术平均值, 其相对标准不确定度 $u(x_1)/x_1 = 0.25\%, u(x_2)/x_2 = 0.57\%, u(x_3)/x_3 = 0.82\%$ 。在这种情况下, $c_i = \partial f / \partial X_i = Y/X_i$ (在 x_1, x_2, x_3 处评估, 见 5.1.3 注 1), $[u_c(y)/y]^2 = \sum_{i=1}^3 [u(x_i)/x_i]^2 = (1.03\%)^2$ (见 5.1.6 注 2), 且式(E.2b)可转换为:

$$\nu_{\text{eff}} = \frac{[u_c(y)/y]^4}{\sum_{i=1}^3 \frac{[u(x_i)/x_i]^4}{\nu_i}}$$

因此

$$\nu_{\text{eff}} = \frac{1.03^4}{\frac{0.25^4}{10-1} + \frac{0.57^4}{5-1} + \frac{0.82^4}{15-1}} = 19.0$$

对 $p = 95\%$ 和 $\nu = 19$ 的 t_p 值, 查表 E.2 表得 $t_{95}(19) = 2.09$; 因此该包含概率的相对扩展不确定度为 $U_{95} = 2.09 \times (1.03\%) = 2.2\%$, 于是 $Y = y \pm U_{95} = y(1 \pm 0.022)$ (y 由 $y = bx_1 x_2 x_3$ 确定), 或写成 $0.978y \leq Y \leq 1.022y$, 及该区间的包含概率约为 95%。

E.4.2 实际中, $u_c(y)$ 取决于输入量的估计值的标准不确定度 $u(x_i)$, 包括正态分布和非正态分布的输入量, $u(x_i)$ 是由基于频率的分布和先验分布得到的 (也就是既可由 A 类评定也可由 B 类评定得到)。对于估计值 y 和与 y 有关的输入量的估计值 x_i 都要加类似的说明。然而, 函数 $t = (y - Y)/u_c(y)$ 的概率分布可用 t 分布近似, 如果它是在期望值附近按泰勒级数展开。大体上, 在最低阶近似中用韦尔奇—萨特思韦特公式得到的就是式(E.2a)或式(E.2b)。

当用式(E.2b)计算 ν_{eff} 时, 提出了如何确定 B 类评定得到的标准不确定度的自由度的问题。根据自由度的定义认识到, t 分布中出现的 ν 是方差 $s^2(\bar{z})$ 的不确定度的度量, C.4.3 中的式(C.7)可用于定义自由度 ν_i 。

$$\nu_i \approx \frac{1}{2} \frac{u^2(x_i)}{2\sigma^2[u(x_i)]} \approx \frac{1}{2} \left[\frac{\Delta u(x_i)}{u(x_i)} \right]^{-2} \quad \dots\dots\dots (E.3)$$

式(E.3)方括号中的量是 $u(x_i)$ 的相对不确定度; 对 B 类评定的标准不确定度, 它是一个主观量, 其值是根据综合可得到的信息进行科学判断来得到的。

示例:

如果对输入量的估计值 x_i 及其标准不确定度 $u(x_i)$ 评定的知识导致判断的 $u(x_i)$ 的值有 25% 的可靠性, 这就意味着相对不确定度取为 $\Delta u(x_i)/u(x_i) = 0.25$, 因此由式(E.3)得 $\nu_i = (0.25)^{-2}/2 = 8$ 。如果判断 $u(x_i)$ 的值大约有 50% 的

可靠性,则 $\nu_i = 2$ (见附录 C 中的表 C.1)。

E.4.3 在 4.3 和 4.4 中关于由先验概率分布评定 B 类标准不确定度的讨论中,其实假设了由这样评定得到的 $u(x_i)$ 值是准确知道的。例如,在 4.3.7 和 4.4.5 中,由假设半宽度 $a = (a_+ - a_-)/2$ 的矩形分布得到 $u(x_i)$ 时, $u(x_i) = a/\sqrt{3}$ 被看作为一个没有不确定度的常数,因为 a_+ 和 a_- 及这样得到的 a 是被看作没有不确定度的(但是见 4.3.9,注 2)。这意味着由式(E.3)看作 $\nu_i \rightarrow \infty$ 或 $1/\nu_i \rightarrow 0$,但它在评定公式(E.2b)中不会引起麻烦。进一步说,假设 $\nu_i \rightarrow \infty$ 不一定是现实的,通常实际上选择 a_- 和 a_+ 时使被测量值落在 a_- 到 a_+ 区间外的概率是极小的。

E.5 其他考虑

E.5.1 在测量不确定度的文章中找到一种表达式,它常被用于获得具有 95% 包含概率的区间,该式为:

$$U'_{95} = [t_{95}^2(\nu'_{\text{eff}})s^2 + 3u^2]^{1/2} \dots\dots\dots (E.4)$$

这里 $t_{95}(\nu'_{\text{eff}})$ 取自 t 分布在自由度为 ν'_{eff} 和 $p = 95\%$ 时的值; ν'_{eff} 是由韦尔奇-萨特思韦特公式[式(E.2b)]计算得到的有效自由度,仅考虑在本次测量中由重复观测值统计评定得到的那些标准不确定度分量 s_i ; $s^2 = \sum c_i^2 s_i^2$; $c_i = \partial f / \partial \theta x_i$; 并且 $u^2 = \sum u_j^2(y) = \sum c_j^2 (a_j^2/3)$ 考虑了不确定度的所有其他分量,其中 $+a_j$ 和 $-a_j$ 假设为 X_j 相对于最佳估计值 x_j 的准确已知的上限和下限(即 $x_j - a_j \leq X_j \leq x_j + a_j$)。

注:本次测量外进行重复观测得到的分量也如同 u^2 中包括的其他分量一样处理。因此为了对式(E.4)和下一节中式(E.5)进行有意义的比较,假设如果存在这种分量则是可以忽略的。

E.5.2 如果提供了包含概率为 95% 的区间的扩展不确定度是根据式 E.3 和 E.4 推荐的方法评定的,可用下式代替式(E.4):

$$U_{95} = t_{95}(\nu_{\text{eff}})(s^2 + u^2)^{1/2} \dots\dots\dots (E.5)$$

其中, ν_{eff} 是根据式(E.2b)计算得到的,并且该计算包括了所有的不确定度分量。

在大多数情况下,由式(E.5)得到的 U_{95} 值比由式(E.4)得到的 U'_{95} 值大,如果假设在用式(E.5)评定时,所有的 B 类方差是由先验矩形分布得到的,该矩形分布的半宽度与用于式(E.4)计算 u^2 的上下限 a_j 一样。这就可以理解,尽管在大多数的情况下 $t_{95}(\nu'_{\text{eff}})$ 稍大于 $t_{95}(\nu_{\text{eff}})$,这两个因子都接近于 2。在式(E.5)中 u^2 是乘以 $t_p^2(\nu_{\text{eff}}) \approx 4$,而在式(E.4)中是乘以 3。虽然 $u^2 \ll s^2$ 时,两式产生相等的值和 U'_{95} ,如果 $u^2 \gg s^2$ 则 U'_{95} 会比 U_{95} 小 13%。因此,一般情况下式(E.4)产生的不确定度提供的区间具有的包含概率比由式(E.5)计算的扩展不确定度提供的区间的包含概率小。

注 1: 在极限情况下 $u^2/s^2 \rightarrow \infty$ 且 $\nu_{\text{eff}} \rightarrow \infty$, $U'_{95} \rightarrow 1.732u$, 而 $U_{95} \rightarrow 1.960u$ 。在这种情况下, U'_{95} 提供了一个具有包含概率为 91.7% 的区间,而 U_{95} 提供了一个具有包含概率为 95% 的区间。当由上限和下限的估计值获得的分量是主要的,且数量多,并且数值大小与 $u_j^2(y) = c_j^2 a_j^2/3$ 相当,这种情况实际上是近似的。

注 2: 对于正态分布,包含因子 $k = \sqrt{3} \approx 1.732$ 提供了包含概率 $p = 91.763\%$ 的区间,这个 p 值与任何其他值相比在一定意义上说是可靠的,也就是对于输入量稍偏离正态的最佳独立。

E.5.3 个别情况,输入量 X_i 是非对称分布的(见 4.3.8)。虽然这在评定 X_i 的估计值 $u(x_i)$ 的标准不确定度 $u_c(y)$ 时没有什么区别,在评定 $u_c(y)$ 中也没有什么区别,但它可能会影响 U 的计算。

通常是给出一个对称区间 $Y = y \pm U$,除非区间内正负方向的偏差不同。如果 X_i 的不对称,仅导致由测量结果 y 及其合成标准不确定度 $u_c(y)$ 表征的概率分布出现小的不对称,则因采用对称区间而导致的一侧的概率损失可以由另一侧增加的概率来补偿。另一种方法是给出一个概率对称的区间(因此 U 不对称):使 Y 落在低于下限 $y - U_-$ 的概率等于 Y 落在高于上限 $y + U_+$ 的概率。但为了引用这样的极限,与简单估计 y 和 $u_c(y)$ 相比需要更多的信息[因此需要每一个输入量 X_i 的简单估计 x_i 和 $u(x_i)$ 外更多的信息]。

E.5.4 这里以 $u_c(y)$ 、 ν_{eff} 和 t 分布的因子 $t_p(\nu_{\text{eff}})$ 评定的扩展不确定 U_p 仅是一种近似,它有局限性。只有在 Y 的分布为正态分布,估计值 y 及其合成标准不确定度 $u_c(y)$ 都是独立的,且假设 $u_c^2(y)$ 的分布是 χ^2 分布时, $(y - Y)/u_c(y)$ 的分布才为 t 分布。式(E.2b)引出的 ν_{eff} 仅涉及了后一个问题,为 $u_c^2(y)$ 提供

了一个近似的 χ^2 分布；另一部分的问题是，要解决由 Y 分布为非正态分布引起的问题，还应考虑除方差外的更高阶矩。

E.6 总结和结论

E.6.1 包含因子 k_p 提供了一个具有接近于规定值的包含概率 p 的区间，如果对每个输入量的概率分布有广泛的了解，并由这些量的分布合成得到的输出量的分布，就可能找到 k_p 的值。这里仅有输入量的估计值 x_i 及其标准不确定度 $u(x_i)$ 是不够的。

E.6.2 由于合成概率分布需要大量的计算，所以很难由可获得的信息的范围和可靠性来证实，故采用输出量的近似分布也是可以接受的。根据中心极限定理，通常假设 $(y - Y)/u_c(y)$ 的概率分布为 t 分布，取 $k_p = t_p(\nu_{\text{eff}})$ ， ν_{eff} 由式 (E.2b) 的韦尔奇-萨特思韦特公式得到，根据 $u_c(y)$ 的有效自由度 ν_{eff} 得到 t 因子。

E.6.3 为了从式 (E.2b) 得到 ν_{eff} ，要求得到每个标准不确定度分量的自由度 ν_i 。对于一个由 A 类评定得到的分量， ν_i 是根据相应输入量独立重复观测值的次数和由这些观测值确定的独立量的数目获得的（见 E.3.3）。对于一个由 B 类评定得到的分量， ν_i 是由该分量值判断出的可靠性得到的 [见 E.4.2 和式 (E.3)]。

E.6.4 计算扩展不确定度优选的方法总结如下，扩展不确定度 $U_p = k_p u_c(y)$ 用于提供一个包含概率约为 p 的区间 $Y = y \pm U_p$ ：

- 1) 按第 4 章和第 5 章所述得到 y 和 $u_c(y)$ 。
- 2) 根据式 (E.2b) 的韦尔奇-萨特思韦特公式计算 ν_{eff} (为方便参考，这里重复给出)：

$$\nu_{\text{eff}} = \frac{u_c^4(y)}{\sum_{i=1}^N \frac{u_i^4(y)}{\nu_i}} \dots\dots\dots (E.2b)$$

如果 $u(x_i)$ 是由 A 类评定得到的，按 E.3.3 提出的要点确定 ν_i 。如果 $u(x_i)$ 是由 B 类评定得到并作为准确已知的来处理，实际情况通常如此， $\nu_i \rightarrow \infty$ ；否则，由式 (E.3) 估计 ν_i 。

- 3) 根据所要求的包含概率 p ，从表 E.2 查得 t 分布的因子 $t_p(\nu_{\text{eff}})$ 。如果 ν_{eff} 不是整数，可将 ν_{eff} 内插或修约到靠近的较小的整数。
- 4) 取 $k_p = t_p(\nu_{\text{eff}})$ ，并计算 $U_p = k_p u_c(y)$ 。

E.6.5 还有一些在实际中很少见的情况，中心极限定理要求的条件不能得到很好地满足，使用 E.6.4 的方法可能会导致不可接受的结果。例如，如果 $u_c(y)$ 中由矩形分布评定的不确定度分量占优势，其界限假设为准确已知，如果 $t_p(\nu_{\text{eff}}) > \sqrt{3}$ 由 U_p 定义的区域上限和下限为 $y + U_p$ 和 $y - U_p$ ，可能落在输出量 Y 的概率分布的界限之外。这种情况应单独处理，但通常可用近似的解析处理（例如，作正态分布与一个矩形分布的卷积^[10]）。

E.6.6 对于在广阔的领域中进行的许多实际测量，通常具备以下条件：

- 被测值 Y 的估计值 y 是由适当个输入量 X_i 的估计值 x_i 得到的， X_i 可用概率分布很好地描述，例如正态分布和矩形分布；
- 这些估计值的标准不确定度 $u(x_i)$ ，可以由 A 类或 B 类评定，对测量结果 y 的合成标准不确定度 $u_c(y)$ 贡献了大小相当的量；
- 不确定度传播率的线性近似是适当的（见 5.1.2 和 C.3.1）；
- $u_c(y)$ 的不确定度很小，因为它的有效自由度 ν_{eff} 具有足够大的值，如大于 10。

在这些情况下，因为符合中心极限定理，由测量结果及其合成标准不确定度表征的概率分布可以假设为正态分布；因为 ν_{eff} 足够大， $u_c(y)$ 可以作为正态分布的标准差的可靠估计值。于是，根据本附录中的讨论，包括强调不确定度评定过程的近似性和区分包含概率相差 1% 或 2% 的区间的不实际性，可以

做下列工作：

——取 $k=2$ 并假设 $U=2u_c(y)$ 确定一个包含概率为近似 95% 的区间；

或，在更严格的应用时，

——取 $k=3$ 并假设 $U=3u_c(y)$ 确定一个包含概率为近似 99% 的区间。

虽然这种方法应该适用于很多实际测量，它对任意特定测量的适用性将取决于 $k=2$ 时与 $t_{95}(\nu_{\text{eff}})$ 的接近程度或 $k=3$ 时与 $t_{99}(\nu_{\text{eff}})$ 的接近程度；也就是说，由 $U=2u_c(y)$ 或 $U=3u_c(y)$ 确定的区间的包含概率分别接近 95% 或 99% 的程度。虽然对 $\nu_{\text{eff}}=11$ ， $k=2$ 和 $k=3$ 会低估 $t_{95}(11)$ 和 $t_{99}(11)$ 约 10% 和 4%（见表 E.2），这在某些情况下是不能接受的。进而，对所有的稍大于 13 的 ν_{eff} 值， $k=3$ 产生了一个包含概率大于 99% 的区间（见表 E.2，该表也表明了对 $\nu_{\text{eff}} \rightarrow \infty$ ，由 $k=2$ 和 $k=3$ 产生的区间的包含概率分别为 95.45% 和 99.73%）。因此，实际上， ν_{eff} 的大小和对扩展不确定度的要求将决定该方法是否可用。

表 E.2 t 分布的 $t_p(\nu)$ 值表
 （自由度为 ν ，从 $-t_p(\nu)$ 到 $+t_p(\nu)$ 的区间包含了分布的分数 p ）

自由度 ν	$p/\%$					
	68.27 ^{a)}	90	95	95.45 ^{a)}	99	99.73 ^{a)}
1	1.84	6.31	12.71	13.97	63.66	235.80
2	1.32	2.92	4.30	4.53	9.92	19.21
3	1.20	2.35	3.18	3.31	5.84	9.22
4	1.14	2.13	2.78	2.87	4.60	6.62
5	1.11	2.02	2.57	2.65	4.03	5.51
6	1.09	1.94	2.45	2.52	3.71	4.90
7	1.08	1.89	2.36	2.43	3.50	4.53
8	1.07	1.86	2.31	2.37	3.36	4.28
9	1.06	1.83	2.26	2.32	3.25	4.09
10	1.05	1.81	2.23	2.28	3.17	3.96
11	1.05	1.80	2.20	2.25	3.11	3.85
12	1.04	1.78	2.18	2.23	3.05	3.76
13	1.04	1.77	2.16	2.21	3.01	3.69
14	1.04	1.76	2.14	2.20	2.98	3.64
15	1.03	1.75	2.13	2.18	2.95	3.59
16	1.03	1.75	2.12	2.17	2.92	3.54
17	1.03	1.74	2.11	2.16	2.90	3.51
18	1.03	1.73	2.10	2.15	2.88	3.48
19	1.03	1.73	2.09	2.14	2.86	3.45
20	1.03	1.72	2.09	2.13	2.85	3.42
25	1.02	1.71	2.06	2.11	2.79	3.33
30	1.02	1.70	2.04	2.09	2.75	3.27
35	1.01	1.70	2.03	2.07	2.72	3.23

表 E.2 (续)

自由度 ν	$p/\%$					
	68.27 ^{a)}	90	95	95.45 ^{a)}	99	99.73 ^{a)}
40	1.01	1.68	2.02	2.06	2.70	3.20
45	1.01	1.68	2.01	2.06	2.69	3.18
50	1.01	1.68	2.01	2.05	2.68	3.16
100	1.005	1.66	1.984	2.025	2.626	3.077
∞	1.000	1.645	1.960	2.000	2.576	3.000

^{a)} 对某量 z 可用数学期望为 μ_z 、标准差 σ 的正态分布描述; 区间为 $\mu_z \pm k\sigma$, 当 $k=1, 2$ 和 3 时, 该区间包含分布的分数 $p=68.27\%$ 、 95.45% 和 99.73% 。

附 录 F
(资料性附录)
示 例

本附录给出了 F.1 至 F.6 共 6 个示例,其目的是对本标准中关于评定与表示测量不确定度的基本原理给出详细的说明。本附录以及本标准的正文和其他附录中的示例都是用于指导用户将这些原理应用到他们的工作中。

由于举例的目的是为了说明问题,所以做了必要的简化。此外,例子和数据都是用来说明本标准的原理的,因此不应当理解为是实际测量的描述。为了避免舍入误差,这里给出的中间过程计算用的数据比通常保留了更多的位数,这样,所述的若干量的计算结果可能与本文中给出的这些量的数值结果稍有不同。

本标准的前面部分指出,将不确定度分量的评定方法分为 A 类评定和 B 类评定,仅仅是因为方便,对测量结果的合成标准不确定度或扩展不确定度的确定,就不需要分类了,因为所有不确定度分量不论是按哪种方法评定的,处理方法是一样的(参见 B.1.3.4, 5.1.2 和 C.3.7)。因此,在举例中不专门去识别特定的不确定度分量的评定方法属于哪一类。然而,从讨论中就会很清楚该分量由 A 类评定还是 B 类评定得到。

F.1 量块的校准

这个例子说明,即使一个看起来很简单的测量,也包含了不确定度评定的细节。

F.1.1 测量问题

一个标称值 50 mm 的量块,其长度是通过比较它和已知标准(具备相同的标称长度)的长度得到的。两个量块比较的直接输出是它们的长度差:

$$d = l(1 + \alpha\theta) - l_s(1 + \alpha_s\theta_s) \dots\dots\dots(F.1)$$

- 其中, l 为被测量,即被校量块在 20 °C 时的长度;
- l_s 为校准证书上给出的 20 °C 条件下标准量块的长度;
- α 和 α_s 分别为被校准量块和标准量块的热膨胀系数;
- θ 和 θ_s 分别为被校准量块和标准量块温度相对于 20 °C 参考温度的偏差。

F.1.2 数学模型

由(F.1)可知,被测量由下式给出:

$$l = \frac{l_s(1 + \alpha_s\theta_s) + d}{(1 + \alpha\theta)} = l_s + d + l_s(\alpha_s\theta_s - \alpha\theta) + \dots \dots\dots(F.2)$$

如果把被校准量块和标准量块的温差表示为 $\delta\theta = \theta - \theta_s$,热膨胀系数差表示为 $\delta\alpha = \alpha - \alpha_s$,式(F.2)变为

$$l = f(l_s, d, \alpha_s, \theta, \delta\alpha, \delta\theta) = l_s + d - l_s(\delta\alpha \cdot \theta + \alpha_s \cdot \delta\theta) \dots\dots\dots(F.3)$$

差值 $\delta\theta$ 和 $\delta\alpha$,估计值为 0,但它们的不确定度不为零;假设 $\delta\alpha, \alpha_s, \delta\theta$ 和 θ 不相关(如果被测量用变量 θ, θ_s, α 和 α_s 表示,就需要考虑 θ 与 θ_s 间及 α 与 α_s 间的相关性)。

由式(F.3)可知,被测量 l 的估计值可以简单地表达为 $l_s + \bar{d}$,其中 l_s 是校准证书上给出的 20 °C 条件下的标准量块长度, \bar{d} 是 d 的估计值,是 $n = 5$ 次独立重复观测的算术平均值。由 5.1.2 的式(10)和式(F.3)可以得到 l 的合成标准不确定度 $u_c(l)$,下面将进行讨论。

注：在该例和其他例子中，为简单起见，一个量与量的估计值用同样的符号。

F.1.3 有贡献的方差

此处及后文中，对本例讨论的有关方法，总结在表 F.1 中。

既然假设 $\delta\alpha=0, \delta\theta=0$ ，将 5.1.2 的式(10)用到式(F.3)上，可以得到：

$$u_c^2(l) = c_s^2 u^2(l_s) + c_d^2 u^2(d) + c_{\alpha_s}^2 u^2(\alpha_s) + c_\theta^2 u^2(\theta) + c_{\delta\alpha}^2 u^2(\delta\alpha) + c_{\delta\theta}^2 u^2(\delta\theta) \dots\dots (F.4)$$

其中，

$$c_s = \partial f / \partial l_s = 1 - (\delta\alpha \cdot \theta + \alpha_s \cdot \delta\theta) = 1$$

$$c_d = \partial f / \partial d = 1$$

$$c_{\alpha_s} = \partial f / \partial \alpha_s = -l_s \delta\theta = 0$$

$$c_\theta = \partial f / \partial \theta = -l_s \delta\alpha = 0$$

$$c_{\delta\alpha} = \partial f / \partial \delta\alpha = -l_s \theta$$

$$c_{\delta\theta} = \partial f / \partial \delta\theta = -l_s \alpha_s$$

因此

$$u_c^2(l) = u^2(l_s) + u^2(d) + l_s^2 \theta^2 u^2(\delta\alpha) + l_s^2 \alpha_s^2 u^2(\delta\theta) \dots\dots\dots (F.5)$$

F.1.3.1 标准量块校准引入的不确定度 $u(l_s)$

校准证书给出了标准量块的扩展不确定度 $U=0.075 \mu\text{m}$ ，包含因子 $k=3$ 。那么标准不确定度为：

$$u(l_s) = (0.075 \mu\text{m}) / 3 = 25 \text{ nm}$$

F.1.3.2 测得的长度差引入的不确定度 $u(d)$

用对两个标准量块的长度差进行 25 次独立重复观测的变异性来确定表征 l 和 l_s 比较时的合并实验标准差，得到其值为 13 nm。在本次比较时取 5 次重复观测，则 5 次读数的算术平均值的标准不确定度为(见 4.2.4)：

$$u(\bar{d}) = s(\bar{d}) = (13 \text{ nm}) / \sqrt{5} = 5.8 \text{ nm}$$

根据比较 l 和 l_s 所用的比较仪的校准证书，由于“随机误差”引起的不确定度为 $\pm 0.01 \mu\text{m}$ ，其包含概率为 95%，并由 6 次重复测量得到；这样，对于自由度 $\nu=6-1=5$ ，使用 t 因子 $t_{95}(5)=2.57$ (参见附录 E，表 E.2)，得到标准不确定度为：

$$u(d_1) = (0.01 \mu\text{m}) / 2.57 = 3.9 \text{ nm}$$

校准证书上给出的由于“系统误差”引起的比较仪的不确定度按 3 倍标准差计为 $0.02 \mu\text{m}$ ，由此引入的标准不确定度为：

$$u(d_2) = (0.02 \mu\text{m}) / 3 = 6.7 \text{ nm}$$

由估计方差的和得到总的贡献：

$$u^2(d) = u^2(\bar{d}) + u^2(d_1) + u^2(d_2) = 93 \text{ nm}^2, \\ \text{或 } u(d) = 9.7 \text{ nm}$$

F.1.3.3 热膨胀系数引入的不确定度 $u(\alpha_s)$

标准量块的热膨胀系数为 $\alpha_s = 11.5 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ ，其不确定性呈矩形分布，区间为 $\pm 2 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ ，则标准不确定度为[见 4.3.7 中式(7)]：

$$u(\alpha_s) = (2 \times 10^{-6} \text{ }^\circ\text{C}^{-1}) / \sqrt{3} = 1.2 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$$

由 F.1.3 可知 $c_{\alpha_s} = \partial f / \partial \alpha_s = -l_s \delta\theta = 0$ ，说明该量对 l 的不确定度不引入一阶贡献，然而它具有二阶贡献，关于这方面的问题在 F.1.7 中讨论。

F.1.3.4 量块温度的偏差引入的不确定度

报告给出的测试台温度为 $(19.9 \pm 0.5)^\circ\text{C}$ ；单次观测时的温度没有记录，说明温度的最大偏差为 $\Delta = 0.5^\circ\text{C}$ ，也就是说，在热作用系统下温度的近似周期性变化的幅度为 0.5°C ，这个数值不是平均温度的不确定度，平均温度的偏差值为：

$$\bar{\theta} = 19.9^\circ\text{C} - 20^\circ\text{C} = -0.1^\circ\text{C}$$

由于测试台的平均温度的不确定性引起的 $\bar{\theta}$ 的标准不确定度为：

$$u(\bar{\theta}) = 0.2^\circ\text{C}$$

当温度随时间周期变化，将产生温度的U形(反正弦)分布，引入的标准不确定度为：

$$u(\Delta) = (0.5^\circ\text{C})/\sqrt{2} = 0.35^\circ\text{C}$$

温度偏差 θ 可取为等于 $\bar{\theta}$ ，则 θ 的标准不确定度为：

$$u^2(\theta) = u^2(\bar{\theta}) + u^2(\Delta) = 0.165^\circ\text{C}^2$$

得到， $u(\theta) = 0.41^\circ\text{C}$

由F.1.3可知 $c_\theta = \partial f / \partial \theta = -l_s \delta \alpha = 0$ ，这一不确定度也对 l 的不确定度不引入一阶贡献，然而它具有二阶贡献，见F.1.7。

表 F.1 标准不确定度分量汇总表

标准不确定度分量 $u(x_i)$	不确定度来源	标准不确定度值 $u(x_i)$	$c_i \equiv \partial f / \partial x$	$u_i(l) \equiv c_i u(x_i)$ (nm)	自由度
$u(l_s)$	标准量块的校准	25 nm	1	25	18
$u(d)$	测得的量块间的差值	9.7 nm	1	9.7	25.6
$u(\bar{d})$	重复观测	5.8 nm			24
$u(d_1)$	比较仪的随机影响	3.9 nm			5
$u(d_2)$	比较仪的系统影响	6.7 nm			8
$u(\alpha_s)$	标准量块的热膨胀系数	$1.2 \times 10^{-6}^\circ\text{C}^{-1}$	0	0	
$u(\theta)$	测试台的温度	0.41°C	0	0	
$u(\bar{\theta})$	测试台的平均温度	0.2°C			
$u(\Delta)$	室温的周期变化	0.35°C			
$u(\delta\alpha)$	量块膨胀系数的差异	$0.58 \times 10^{-6}^\circ\text{C}^{-1}$	$-l_s \theta$	2.9	50
$u(\delta\theta)$	量块的温度差异	0.029°C	$-l_s \alpha_s$	16.6	2
				$u_c^2(l) = \sum u_i^2(l) = 1\,002\text{ nm}^2$ $u_c(l) = 32\text{ nm}$ $\nu_{\text{eff}}(l) = 16$	

F.1.3.5 膨胀系数差异引入的不确定度 $u(\delta\alpha)$

$\delta\alpha$ 的变异性估计限为 $\pm 1 \times 10^{-6}^\circ\text{C}^{-1}$ ， $\delta\alpha$ 的任意值以等概率落在此范围内，其标准不确定度为

$$u(\delta\alpha) = (1 \times 10^{-6}^\circ\text{C}^{-1})/\sqrt{3} = 0.58 \times 10^{-6}^\circ\text{C}^{-1}$$

F.1.3.6 量块的温度差引入的不确定度 $u(\delta\theta)$

标准量块和被校准量块被认为是在同一温度下，但实际上存在着温度差，温度差以等概率落在估计

区间 $-0.05\text{ }^{\circ}\text{C}$ 到 $0.05\text{ }^{\circ}\text{C}$ 内,则标准不确定度为:

$$u(\delta\theta) = (0.05\text{ }^{\circ}\text{C})/\sqrt{3} = 0.029\text{ }^{\circ}\text{C}$$

F.1.4 合成标准不确定度

由式(F.5)可以计算合成标准不确定度 $u_c(l)$,代入所有分项可以得到:

$$u_c^2(l) = (25\text{ nm})^2 + (9.7\text{ nm})^2 + (0.05\text{ m})^2(-0.1\text{ }^{\circ}\text{C})^2(0.58 \times 10^{-6}\text{ }^{\circ}\text{C}^{-1})^2 + \\ (0.05\text{ m})^2(11.5 \times 10^{-6}\text{ }^{\circ}\text{C}^{-1})^2(0.029\text{ }^{\circ}\text{C})^2 \dots\dots\dots(\text{F.6a})$$

$$= (25\text{ nm})^2 + (9.7\text{ nm})^2 + (2.9\text{ nm})^2 + (16.6\text{ nm})^2 = 1\,002\text{ nm}^2 \dots\dots\dots(\text{F.6b})$$

或

$$u_c(l) = 32\text{ nm} \dots\dots\dots(\text{F.6c})$$

很明显,不确定度的主要分量是标准量块的不确定度 $u(l_s) = 25\text{ nm}$ 。

F.1.5 最终结果

校准证书给出了标准量块 $20\text{ }^{\circ}\text{C}$ 时的长度为 $l_s = 50.000\,623\text{ mm}$ 。被测量块和标准量块的长度差的5次重复观测结果的算术平均值 \bar{d} 为 215 nm 。这样,由F.1.2有 $l = l_s + \bar{d}$,则被测量块的长度在 $20\text{ }^{\circ}\text{C}$ 时为 $50.000\,838\text{ mm}$ 。根据7.2.2,测量的最终结果可以表示为: $l = 50.000\,838\text{ mm}$,合成标准不确定度 $u_c = 32\text{ nm}$ 。相应的相对合成标准不确定度为 $u_c/l = 6.4 \times 10^{-7}$ 。

F.1.6 扩展不确定度

假设需要得到扩展不确定度 $U_{99} = k_{99} u_c(l)$,它提供一个包含概率近似为99%的区间。使用的步骤总结在E.6.4中,自由度见表F.1,过程如下:

- 1) 标准量块校准引入的不确定度, $u(l_s)$ [F.1.3.1]。校准证书说明,计算扩展不确定度所引用的合成标准不确定度的有效自由度为 $\nu_{\text{eff}}(l_s) = 18$ 。
- 2) 测得的长度差引入的不确定度, $u(d)$ [F.1.3.2]。尽管 \bar{d} 由5次重复测量得到,因为 $u(\bar{d})$ 是根据25次观测的合并实验标准差得到的,所以, $u(\bar{d})$ 的自由度为 $\nu(\bar{d}) = 25 - 1 = 24$ (见F.3.6,注)。比较仪由于随机影响引起的不确定度 $u(d_1)$ 的自由度 $\nu(d_1) = 6 - 1 = 5$,因为 d_1 是由6次重复测量得到。比较仪由系统影响引入的不确定度 $\pm 0.02\text{ }\mu\text{m}$ 可以假设为可靠到25%,因此由E.4.2的(E.3)得到自由度为 $\nu(d_2) = 8$ (参见E.4.2的示例)。 $u(d)$ 的有效自由度 $\nu_{\text{eff}}(d)$,可由E.4.1的式(E.2b)得到:

$$\nu_{\text{eff}}(d) = \frac{[u^2(\bar{d}) + u^2(d_1) + u^2(d_2)]^2}{\frac{u^4(\bar{d})}{\nu(\bar{d})} + \frac{u^4(d_1)}{\nu(d_1)} + \frac{u^4(d_2)}{\nu(d_2)}} = \frac{(9.7\text{ nm})^4}{\frac{(5.8\text{ nm})^4}{24} + \frac{(3.9\text{ nm})^4}{5} + \frac{(6.7\text{ nm})^4}{8}} = 25.6$$

- 3) 膨胀系数差值引入的不确定度, $u(\delta\alpha)$ [F.1.3.5]。 $\delta\alpha$ 的变异性的估计限为 $\pm 1 \times 10^{-6}\text{ }^{\circ}\text{C}^{-1}$,它要求可靠到10%。由E.4.2中的式(E.3)可以得到 $\nu(\delta\alpha) = 50$ 。
- 4) 量块温度差引入的不确定度, $u(\delta\theta)$ [F.1.3.6]。温度差异 $\delta\theta$ 的估计区间为 $-0.05\text{ }^{\circ}\text{C}$ 到 $+0.05\text{ }^{\circ}\text{C}$,相信可靠到50%,由E.4.2中的式(E.3)可以得到 $\nu(\delta\theta) = 2$ 。

由E.4.1中的式(E.2b)的步骤计算 $\nu_{\text{eff}}(l)$ 的方法与上面2)中计算 $\nu_{\text{eff}}(d)$ 的方法完全一样。因此,由(F.6b)和(F.6c)及由1)到4)给出的自由度值可得到:

$$\nu_{\text{eff}}(l) = \frac{(32\text{ nm})^4}{\frac{(25\text{ nm})^4}{18} + \frac{(9.7\text{ nm})^4}{25.6} + \frac{(2.9\text{ nm})^4}{50} + \frac{(16.6\text{ nm})^4}{2}} = 16.7$$

为了得到所需的扩展不确定度,该值截到相近的最小整数,即 $\nu_{\text{eff}}(l)=16$,根据附录 E 中的表 E.2 可知, $t_{99}(16)=2.92$,因此 $U_{99}=t_{99}(16)u_c(l)=2.92 \times (32 \text{ nm})=93 \text{ nm}$ 。根据 7.2.4,最后的测量结果可以表述为:

$$l=(50.000\ 838 \pm 0.000\ 093) \text{ mm}$$

其中“±”号后的数字为扩展不确定度 $U=ku_c$, U 由合成标准不确定度 $u_c=32 \text{ nm}$ 和包含因子 $k=2.92$ (基于 t 分布,自由度 $\nu=16$) 来确定。该扩展不确定度规定了一个包含概率为 99% 的估计区间。相应的相对扩展不确定度为 $U/l=1.9 \times 10^{-6}$ 。

F.1.7 二阶项

5.1.2 的注指出,用于得到合成标准不确定度 $u_c(l)$ 的式(10),如果函数 $Y=f(X_1, X_2 \dots X_N)$ 的非线性足够明显,使得泰勒展开式中的高次项不能被忽略时,式(10)应扩项。本例正是这种情况,所以在此前所进行的 $u_c(l)$ 的评定是不完全的,将 5.1.2 的注中提供的公式应用到式(F.3),实际上产生了两项明显的、不能忽略的二次项(源自该注的表达式中的平方项),这两项被加进了式(F.5):

$$l_s^2 u^2(\delta\alpha) u^2(\theta) + l_s^2 u^2(\alpha_s) u^2(\delta\theta)$$

但是这些项中只有第一项对 $u_c(l)$ 贡献较大:

$$l_s u(\delta\alpha) u(\theta) = (0.05 \text{ m})(0.58 \times 10^{-6} \text{ }^\circ\text{C}^{-1})(0.41 \text{ }^\circ\text{C}) = 11.7 \text{ nm}$$

$$l_s u(\alpha_s) u(\delta\theta) = (0.05 \text{ m})(1.2 \times 10^{-6} \text{ }^\circ\text{C}^{-1})(0.029 \text{ }^\circ\text{C}) = 1.7 \text{ nm}$$

这个二阶项使 $u_c(l)$ 从 32 nm 增到 34 nm。

F.2 同时测量电阻和电抗

本例说明在同一测量中同时确定多个被测量或多个输出量时的处理方法及它们的估计值的相关性处理。此处只考虑观测值的随机变化;实际上,对系统影响进行修正的修正值的不确定度也对测量结果的不确定度有贡献,对数据用两种不同的方法进行分析,得到基本相同的数值。

F.2.1 测量问题

电路元件的电阻 R 和电抗 X 的测量方法:测量其两端的正弦波交流电位差的幅度 V ,测量通过它的交流电流的幅度 I ,以及交流电位差相对于交流电流的相移角 ϕ 。这样,三个输入量为 V, I, ϕ ,三个输出量 R, X 和 Z ——被测量为三个阻抗成分。因为 $Z^2 = R^2 + X^2$,因此实际上只有两个独立的输出量。

F.2.2 数学模型和数据

根据欧姆定律,输入量和被测量的关系为:

$$R = \frac{V}{I} \cos\phi; X = \frac{V}{I} \sin\phi; Z = \frac{V}{I} \dots\dots\dots (F.7)$$

对输入量 V, I 和 ϕ 的 5 组同时独立的观测值是在类似条件下得到的,数据由表 F.2 给出。根据 4.2 中式(3)和式(5)进行算术平均值和实验标准差的计算,结果也在表中给出。取平均值作为输入量的期望的最佳估计,且把实验标准差作为这些平均值的标准不确定度。

由于平均值 \bar{V}, \bar{I} 和 $\bar{\phi}$ 同时观测得到,它们是相关的,而且其相关性应考虑进被测量 R, X 和 Z 的标准不确定度评定中。所需的相关系数由 5.2.2 的式(14)得到, $s(\bar{V}, \bar{I}), s(\bar{V}, \bar{\phi})$ 和 $s(\bar{I}, \bar{\phi})$ 的值由 5.2.3 的

式(17)计算而来。结果见表 F.2,表中 $r(x_i, x_j) = r(x_j, x_i)$ 且 $r(x_i, x_i) = 1$ 。

表 F.2 由五组同时观测值得到的输入量 V 、 I 和 ϕ 的值

组数 k	输入量		
	V/V	I/mA	ϕ/rad
1	5.007	19.663	1.045 6
2	4.994	19.639	1.043 8
3	5.005	19.640	1.046 8
4	4.990	19.685	1.042 8
5	4.999	19.678	1.043 3
算术平均值	$\bar{V} = 4.999\ 0$	$\bar{I} = 19.661\ 0$	$\bar{\phi} = 1.044\ 46$
均值的实验标准差	$s(\bar{V}) = 0.003\ 2$	$s(\bar{I}) = 0.009\ 5$	$s(\bar{\phi}) = 0.000\ 75$
相关系数			
$r(\bar{V}, \bar{I}) = -0.36$			
$r(\bar{V}, \bar{\phi}) = 0.86$			
$r(\bar{I}, \bar{\phi}) = -0.65$			

F.2.3 结果:方法 1

方法 1 由表 F.3 汇总。

3 个被测量 R 、 X 和 Z 的值用表 F.2 给出的 V 、 I 和 ϕ 的平均值 \bar{V} 、 \bar{I} 和 $\bar{\phi}$ 代入式(F.7)得到, R 、 X 和 Z 的标准不确定度由 5.2.2 的式(16)得到,正如以上指出的,输入量 \bar{V} 、 \bar{I} 和 $\bar{\phi}$ 是相关的。作为一个例子,考虑 $Z = \bar{V}/\bar{I}$ 。用 x_1 表示 \bar{V} ,用 x_2 表示 \bar{I} ,用 $Z = \bar{V}/\bar{I}$ 表示 f ,由 5.2.2 中的式(16)可以得到 Z 的合成标准不确定度:

$$u_c^2(Z) = \left(\frac{1}{\bar{I}}\right)^2 u^2(\bar{V}) + \left(\frac{\bar{V}}{\bar{I}^2}\right)^2 u^2(\bar{I}) + 2\left(\frac{1}{\bar{I}}\right)\left(-\frac{\bar{V}}{\bar{I}^2}\right) u(\bar{V})u(\bar{I})r(\bar{V}, \bar{I}) \dots\dots\dots (F.8a)$$

$$= Z^2 \left[\frac{u(\bar{V})}{\bar{V}}\right]^2 + Z^2 \left[\frac{u(\bar{I})}{\bar{I}}\right]^2 - 2Z^2 \left[\frac{u(\bar{V})}{\bar{V}}\right] \left[\frac{u(\bar{I})}{\bar{I}}\right] r(\bar{V}, \bar{I}) \dots\dots\dots (F.8b)$$

$$\text{或 } u_{c,r}^2(Z) = u_r^2(\bar{V}) + u_r^2(\bar{I}) - 2u_r(\bar{V})u_r(\bar{I})r(\bar{V}, \bar{I}) \dots\dots\dots (F.8c)$$

其中, $u(\bar{V}) = s(\bar{V})$, $u(\bar{I}) = s(\bar{I})$, 最后一个表达式中的 r 表明 u 是相对不确定度。把表 F.2 中的相关数据代入式(F.8a),可以得到 $u_c(Z) = 0.236 \Omega$ 。

因为三个被测量或输出量取决于同样的输入量,所以它们也是相关的。描述这种相关性的协方差矩阵的元素可写成:

$$u(y_l, y_m) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial y_l}{\partial x_i} \frac{\partial y_m}{\partial x_j} u(x_i)u(x_j)r(x_i, x_j) \dots\dots\dots (F.9)$$

其中, $y_l = f_l(x_1, x_2, \dots, x_N)$, $y_m = f_m(x_1, x_2, \dots, x_N)$ 。式(F.9)是 D.1.2.3 中式(D.2)中 q_l 相关时的通用式。输出量的估计相关系数由 $r(y_l, y_m) = u(y_l, y_m)/u(y_l)u(y_m)$ 给出,说明见 5.2.2 的式

(14)。可以证明,协方差矩阵的对角线元素 $u(y_l, y_l) \equiv u^2(y_l)$ 是输出量 y_l 的估计方差(参见 5.2.2 的注 2),当 $m=l$ 时,式(F.9)等同于 5.2.2 的式(16)。

为将式(F.9)用于本例,作以下的标识:

$$y_1 = R, x_1 = \bar{V}, u(x_i) = s(x_i)$$

$$y_2 = X, x_2 = \bar{I}, N = 3$$

$$y_3 = Z, x_3 = \bar{\phi}$$

R 、 X 和 Z 的计算结果,以及它们的估计方差、相关系数由表 F.3 给出。

表 F.3 输出量 R 、 X 和 Z 的计算值:方法 1

被测量序号 l	被测量估计值 y_l 与输入量的估计值 x_i 的关系	测量结果 y_l 的估计值	测量结果的合成标准不确定度 $u_c(y_l)$
1	$y_1 = R = (\bar{V}/\bar{I}) \cos \bar{\phi}$	$y_1 = R = 127.732 \Omega$	$u_c(R) = 0.071 \Omega,$ $u_c(R)/R = 0.06 \times 10^{-2}$
2	$y_2 = X = (\bar{V}/\bar{I}) \sin \bar{\phi}$	$y_2 = X = 219.847 \Omega$	$u_c(X) = 0.295 \Omega,$ $u_c(X)/X = 0.13 \times 10^{-2}$
3	$y_3 = Z = \bar{V}/\bar{I}$	$y_3 = Z = 254.260 \Omega$	$u_c(Z) = 0.236 \Omega,$ $u_c(Z)/Z = 0.09 \times 10^{-2}$
相关系数 $r(y_l, y_m)$			
$r(y_1, y_2) = r(R, X) = -0.588$			
$r(y_1, y_3) = r(R, Z) = -0.485$			
$r(y_2, y_3) = r(X, Z) = -0.993$			

F.2.4 结果:方法 2

方法 2 由表 F.4 汇总。

因为已经获得了 3 个输入量 V 、 I 和 ϕ 的 5 组观测值的数据,可以根据每组输入数据计算出 R 、 X 和 Z 的值,取 5 个值的算术平均值为 R 、 X 和 Z 的最佳估计值。于是,每个平均值的实验标准差(即其合成标准不确定度)可用常用的方法[4.2.3 的式(5)]由 5 个单值计算得到;3 个平均值的估计协方差可用 5.2.3 中式(17)由每个平均值获得的 5 个单值计算得到。除了对二阶影响要用 \bar{V}/\bar{I} 代替 \overline{VI} , $\overline{\cos \phi}$ 代替 $\overline{\cos \phi}$ 以外,两种方法提供的输出值、标准不确定度,以及估计的协方差都没有区别。

为了说明这一方法,表 F.4 给出了由每组观测值计算得到的 R 、 X 和 Z 值。算术平均值、标准不确定度、估计的相关系数都直接由这些值计算得到。该方法得到的数学结果和表 F.3 给出的结果没有什么差别。

表 F.4 输出量 R 、 X 和 Z 的计算值:方法 2

设置值 k	单次测量值		
	$R = (V/I)\cos\phi$ Ω	$X = (V/I)\sin\phi$ Ω	$Z = V/I$ Ω
1	127.67	220.32	254.64
2	127.89	219.79	254.29
3	127.51	220.64	254.84
4	127.71	218.97	253.49
5	127.88	219.51	254.04
算术平均值	$y_1 = \bar{R} = 127.732$	$y_2 = \bar{X} = 219.847$	$y_3 = \bar{Z} = 254.260$
平均值的实验标准差	$s(\bar{R}) = 0.071$	$s(\bar{X}) = 0.295$	$s(\bar{Z}) = 0.236$
相关系数 $r(y_l, y_m)$			
$r(y_1, y_2) = r(\bar{R}, \bar{X}) = -0.588$			
$r(y_1, y_3) = r(\bar{R}, \bar{Z}) = -0.485$			
$r(y_2, y_3) = r(\bar{X}, \bar{Z}) = 0.993$			

在 4.1.4 的注中,方法 2 是由 $\bar{Y} = (\sum_{k=1}^n Y_k)/n$ 获得估计值 y ,而方法 1 是由 $y = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N)$ 获得估计值 y 。就像注中指出的那样,一般情况下,如果 f 是输入量(假设执行方法 1 时已经考虑了实验观测值的相关系数)的线性函数,两种方法将给出同样的结果。如果 f 不是一个线性函数,那么方法 1 的结果将和方法 2 的结果有所不同,不同程度取决于非线性程度、 X_i 的估计方差和协方差,可参考表达式:

$$y = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N) + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2 f}{\partial \bar{X}_i \partial \bar{X}_j} u(\bar{X}_i, \bar{X}_j) + \dots \dots \dots (F.10)$$

其中,公式右边第二项是以项 \bar{X}_i 对 f 的泰勒级数展开中的二阶项(也见于 5.1.2,注)。在目前的情况下,方法 2 更好,因为它避免了 $y = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N)$ 的近似,更好地反映了测量过程——数据实际上是以组为单位采集的。

另一方面,如果表 F.2 中的数据代表了电位差 V 的 $n_1 = 5$ 次观测,电流 I 的 $n_2 = 5$ 次观测和相位 ϕ 的 $n_3 = 5$ 次观测,但如果 $n_1 \neq n_2 \neq n_3$ (事实上,这是很差的测量程序,因为通过一个固定阻抗的电位差和电流是直接相关的),则方法 2 就会不合适。

如果表 F.2 的数据以这种方式重新解释,方法 2 就不合适了,如果量 V 、 I 和 ϕ 之间不相关,或没有明显的相关性,则相关系数应设为 0。此时,式(F.9)就简化为 D.1.2.3 的式(D.2),即:

$$u(y_l, y_m) = \sum_{i=1}^N \frac{\partial y_l}{\partial x_i} \frac{\partial y_m}{\partial x_i} u^2(x_i) \dots \dots \dots (F.11)$$

将表 F.2 的数据带入,引起表 F.3 中的变化示于表 F.5。

表 F.5 假设表 F.2 中相关系数为 0 时表 F.3 的变化

测量结果的合成标准不确定度 $u_c(y_l)$
$u_c(R) = 0.195 \Omega, u_c(R)/R = 0.15 \times 10^{-2}$
$u_c(X) = 0.201 \Omega, u_c(X)/X = 0.09 \times 10^{-2}$
$u_c(Z) = 0.204 \Omega, u_c(Z)/Z = 0.08 \times 10^{-2}$
相关系数 $r(y_l, y_m)$
$r(y_1, y_2) = r(R, X) = 0.056$
$r(y_1, y_3) = r(R, Z) = 0.527$
$r(y_2, y_3) = r(X, Z) = 0.878$

F.3 温度计的校准

该例说明了使用最小二乘法获取线性校准曲线,以及如何用拟合的参数:截距、斜率和它们的估计方差与协方差,由曲线获得修正值及其标准不确定度。

F.3.1 测量问题

温度计是用与已知的参考温度相比较的方法校准的。温度计的温度读数为 t_k ,进行了 11 次读数,每个读数的不确定度可忽略,相应的已知参考温度为 $t_{R,k}$,其温度范围为 21 °C 到 27 °C,由此获得读数的修正量为 $b_k = t_{R,k} - t_k$,测得的修正量 b_k 和测得的温度 t_k 是输入量。线性校准曲线为

$$b(t) = y_1 + y_2(t - t_0) \dots\dots\dots (F.12)$$

是用最小二乘法对测得的修正值和温度拟合成直线。其中,参数 y_1 和 y_2 ,分别代表了校准曲线的截距和斜率,是两个待测定的被测量或输出量。温度 t_0 是所选择的准确的参考温度;用最小二乘法拟合而确定的不是一个独立的参数。一旦找到了 y_1 和 y_2 以及它们的估计方差和协方差,式(F.12)用于预测温度计任意一个温度值 t 的修正值和修正值的标准不确定度。

F.3.2 最小二乘法拟合

根据最小二乘法和 F.3.1 的假设条件,输出量 y_1 和 y_2 以及它们的估计方差、协方差在下式之和为最小时得到。

$$S = \sum_{k=1}^n [b_k - y_1 - y_2(t_k - t_0)]^2$$

这就导出了 y_1 和 y_2 的以下各式,它们的实验方差 $s^2(y_1)$ 和 $s^2(y_2)$,它们的估计的相关系数 $r(y_1, y_2) = s(y_1, y_2) / s(y_1)s(y_2)$,其中 $s(y_1, y_2)$ 是估计的协方差。

$$y_1 = \frac{(\sum b_k)(\sum \theta_k^2) - (\sum b_k \theta_k)(\sum \theta_k)}{D} \dots\dots\dots (F.13a)$$

$$y_2 = \frac{n \sum b_k \theta_k - (\sum b_k)(\sum \theta_k)}{D} \dots\dots\dots (F.13b)$$

$$s^2(y_1) = \frac{s^2 \sum \theta_k^2}{D} \dots\dots\dots (F.13c)$$

$$s^2(y_2) = n \frac{s^2}{D} \dots\dots\dots (F.13d)$$

$$r(y_1, y_2) = -\frac{\sum \theta_k}{\sqrt{n \sum \theta_k^2}} \dots\dots\dots (F.13e)$$

$$s^2 = \frac{\sum [b_k - b(t_k)]^2}{n - 2} \dots\dots\dots (F.13f)$$

$$D = n \sum \theta_k^2 - (\sum \theta_k)^2 = n \sum (\theta_k - \bar{\theta})^2 = n \sum (t_k - \bar{t})^2 \dots\dots\dots (F.13g)$$

其中,所有的求和都是从 $k=1$ 到 n , $\theta_k = t_k - t_0$, $\bar{\theta} = (\sum \theta_k)/n$, 并且 $\bar{t} = (\sum t_k)/n$; $[b_k - b(t_k)]$ 是在 t_k 温度时测得或观测到的修正值 b_k 与拟合曲线 $b(t) = y_1 + y_2(t - t_0)$ 上在 t_k 时预示的修正值 $b(t_k)$ 之间的差值; 方差 s^2 是总的拟合的不确定度的度量, 其中因子 $n - 2$ 反映了这样一个事实, 即由 n 次观测确定两个参数 y_1 和 y_2 , s^2 的自由度为 $\nu = n - 2$ (见 E.3.3)。

F.3.3 结果的计算

被拟合的数据在表 F.6 的第二列和第三列给出, 取 $t_0 = 20$ °C 作为参考温度, 应用(F.13a)到(F.13g)可以得到:

$$y_1 = -0.171\ 2\ \text{°C}, s(y_1) = 0.002\ 9\ \text{°C}$$

$$y_2 = 0.002\ 18, s(y_2) = 0.000\ 67$$

$$r(y_1, y_2) = -0.930, s = 0.003\ 5\ \text{°C}$$

斜率 y_2 比其标准不确定度大了 3 倍还多, 说明需要一个校准曲线, 而不是用一个固定的平均修正值。

表 F.6 用最小二乘法得到温度计线性校准曲线时所用的数据

读数序号 k	温度计的读数 t_k °C	观测的修正值 $b_k = t_{R,k} - t_k$ °C	预计的修正值 $b(t_k)$ °C	观测值与预计的修正值之差 $b_k - b(t_k)$ °C
1	21.521	-0.171	-0.167 9	-0.003 1
2	22.012	-0.169	-0.166 8	-0.002 2
3	22.512	-0.166	-0.165 7	-0.000 3
4	23.003	-0.159	-0.164 6	+0.005 6
5	23.507	-0.164	-0.163 5	-0.000 5
6	23.999	-0.165	-0.162 5	-0.002 5
7	24.513	-0.156	-0.161 4	+0.005 4
8	25.002	-0.157	-0.160 3	+0.003 3
9	25.503	-0.159	-0.159 2	+0.000 2
10	26.010	-0.161	-0.158 1	-0.002 9
11	26.511	-0.160	-0.157 0	-0.003 0

那么, 校准曲线表示为:

$$b(t) = -0.171\ 2(29)\ \text{°C} + 0.002\ 18(67) \times (t - 20\ \text{°C}) \dots\dots\dots (F.14)$$

其中, 括号内的数字是标准不确定度数值。与所说明的截距和斜率的结果的最后位数字相对齐(见 7.2.2), 这一公式给出了修正量 $b(t)$ 在任意温度点 t 上的预计值, 特别是 $t = t_k$ 时的值 $b(t_k)$, 表的第 4 列给出了这些值。最后一列给出了测得值与预计值之间的差值 $b_k - b(t_k)$ 。这些差值的分析可用于

检查线性模型的有效性;有正规的检验方法(见参考文献[8]),但在本例中不作考虑。

F.3.4 预计值的不确定度

修正值的预计值的合成标准不确定度的表达式可由不确定度传播律,将5.2.2中式(16)应用到式(F.12)。注意: $b(t)=f(y_1, y_2)$,且 $u(y_1)=s(y_1)$, $u(y_2)=s(y_2)$,得到:

$$u_c^2[b(t)] = u^2(y_1) + (t - t_0)^2 u^2(y_2) + 2(t - t_0)u(y_1)u(y_2)r(y_1, y_2) \dots\dots (F.15)$$

估计方差 $u_c^2[b(t)]$ 在 $t_{\min} = t_0 - u(y_1)r(y_1, y_2)/u(y_2)$ 时为最小,此时 $t_{\min} = 24.0085\text{ }^\circ\text{C}$ 。作为使用式(F.15)的一个例子,要求在 $t = 30\text{ }^\circ\text{C}$ 时的温度计修正值及其不确定度,这个温度是在温度计实际校准温度的范围外。将 $t = 30\text{ }^\circ\text{C}$ 代入式(F.14)有 $b(30\text{ }^\circ\text{C}) = -0.1494\text{ }^\circ\text{C}$,而式(F.15)变为

$$u_c^2[b(30\text{ }^\circ\text{C})] = (0.0029\text{ }^\circ\text{C})^2 + (10\text{ }^\circ\text{C})^2 \times (0.00067)^2 + 2 \times (10\text{ }^\circ\text{C}) \times (0.0029\text{ }^\circ\text{C}) \times (0.00067) \times (-0.930) \\ = 17.1 \times 10^{-6}\text{ }^\circ\text{C}^2$$

或 $u_c[b(30\text{ }^\circ\text{C})] = 0.0041\text{ }^\circ\text{C}$

这样,30 °C时的修正值为-0.1494 °C,其合成标准不确定度 $u_c = 0.0041\text{ }^\circ\text{C}$,自由度 $\nu = n - 2 = 9$ 。

F.3.5 消除截距和斜率间的相关性

相关系数 $r(y_1, y_2)$ 的式(F.13e)表明:如果 t_0 选得恰到好处,那么 $\sum_{k=1}^n \theta_k = \sum_{k=1}^n (t_k - t_0) = 0$,则 $r(y_1, y_2) = 0$,即 y_1 和 y_2 不相关,从而可以简化预期修正值的标准不确定度的计算。由于当 $t_0 = \bar{t} = (\sum_{k=1}^n t_k)/n$,且 $\bar{t} = 24.0085\text{ }^\circ\text{C}$ 时, $\sum_{k=1}^n \theta_k = 0$,由此,使用 $t_0 = \bar{t} = 24.0085\text{ }^\circ\text{C}$ 进行最小二乘拟合,将使 y_1 和 y_2 不相关(温度 \bar{t} 也是使 $u^2[b(t)]$ 拥有最小值的温度点,见F.3.4)。然而,不必重复拟合,因为:

$$b(t) = y'_1 + y_2(t - \bar{t}) \dots\dots\dots (F.16a)$$

$$u_c^2[b(t)] = u^2(y'_1) + (t - \bar{t})^2 u^2(y_2) \dots\dots\dots (F.16b)$$

$$r(y'_1, y_2) = 0 \dots\dots\dots (F.16c)$$

其中:

$$y'_1 = y_1 + y_2(\bar{t} - t_0)$$

$$\bar{t} = t_0 - s(y_1)r(y_1, y_2)/s(y_2)$$

$$s^2(y'_1) = s^2(y_1)[1 - r^2(y_1, y_2)]$$

式(F.16b)中,做了 $u(y'_1) = s(y'_1)$, $u(y_2) = s(y_2)$ 的替代[参见式(F.15)]。

将这些关系式应用到F.3.3的结果中,可以得到:

$$b(t) = -0.1625(11) + 0.00218(67) \times (t - 24.0085\text{ }^\circ\text{C}) \dots\dots\dots (F.17a)$$

$$u_c^2[b(t)] = (0.0011)^2 + (t - 24.0085\text{ }^\circ\text{C})^2 \times (0.00067)^2 \dots\dots\dots (F.17b)$$

这些表达式给出的结果与式(F.14)和式(F.15)的结果相同,并且可以用 $b(30\text{ }^\circ\text{C})$ 和 $u_c[b(30\text{ }^\circ\text{C})]$ 的计算来核验。将 $t = 30\text{ }^\circ\text{C}$ 代入式(F.17a)和式(F.17b)得到:

$$b(30\text{ }^\circ\text{C}) = -0.1494\text{ }^\circ\text{C}$$

$$u_c[b(30\text{ }^\circ\text{C})] = 0.0041\text{ }^\circ\text{C}$$

与F.3.4结论一致。两个预期修正值 $b(t_1)$ 和 $b(t_2)$ 间的估计协方差可由F.2.3的式(F.9)得到。

F.3.6 其他考虑

最小二乘法可用于数据点的高阶曲线拟合,也可用于具有不确定度的独立数据点。该方面的标准文本可参见参考文献[8]。以下例子说明了被测量的修正值 b_k 假设为不能准确知道时的两种情况。

1) 设每个 t_k 的不确定度可忽略,设 n 个 $t_{R,k}$ 中的每个值都是由 m 次重复观测读数得到,并设读数

的合并估计方差是根据几个月获得的大量数据得到,合并估计方差为 s_p^2 。则每个 $t_{R,k}$ 的估计方差为 $s_p^2/m = u_0^2$,每个观测值的修正值 $b_k = t_{R,k} - t_k$ 具有相同的标准不确定度 u_0 。在这种情况下(假设没有理由认为线性模型是不正确的),在式(F.13c)和式(F.13d)中用 u_0^2 代替 s^2 。

注:样本合并方差 s_p^2 基于对于同一随机变量的 N 次独立观测,由下式得到:

$$s_p^2 = \frac{\sum_{i=1}^N \nu_i s_i^2}{\sum_{i=1}^N \nu_i}$$

其中, S_i^2 是 n_i 中第 i 个独立重复观测序列的实验方差[4.2.2中式(4)],自由度 $\nu_i = n_i - 1$ 。 s_p^2 的自由度是 $\nu = \sum_{i=1}^N \nu_i$ 。用合并估计方差 s_p^2 表征的 m 次独立观测值的算术平均值的实验方差 s_p^2/m (实验标准差 S_p/\sqrt{m}),具有的自由度也为 ν 。

2) 假设每个 t_k 的不确定度可忽略,对 n 个 $t_{R,k}$ 的每个值加修正值 ϵ_k ,且每个修正具有相同的标准不确定度 u_a 。则每个 $b_k = t_{R,k} - t_k$ 的标准不确定度也是 u_a ,并且用 $s^2(y_1) + u_a^2$ 代替了 $s^2(y_1)$,用 $s^2(y'_1) + u_a^2$ 代替 $s^2(y_1)$ 。

F.4 活度的测量

本例与 F.2 中同时测量电阻和电抗的例子相似,其数据可以用两种不同的方法进行分析,但每种方法本质上会得出同样的数值结果。第一种方法再次说明需要考虑观察到的各输入量间的相关性。

F.4.1 测量问题

水样品中未知的氡(^{222}Rn)活度浓度可以用液体闪烁计数方法,通过与一个具有已知活度浓度的含氡水标准样品对比来确定。未知活度浓度通过测量三种计数源得到,这些计数源由大约 5 g 水和 12 g 有机乳胶闪烁体组成,装在 3 个 22 mL 小玻璃瓶中。

- 源(a):质量为 m_s 的活度浓度已知的标准溶液;
- 源(b):不含放射物质的纯水样品,用于获得本底计数率;
- 源(c):质量为 m_x 活度浓度未知的试样构成的样品。

按照标准溶液—净水—样品溶液的顺序,对 3 种计数源进行 6 次循环测量;在所有 6 次循环中,对每种源的每次经过死时间修正的计数时间间隔 T_0 为 60 min。虽然在整个计数区间(65 h)内不能假定本底计数率是不变的,但可以假设每次净水测量所得的计数次数可以作为同一循环中测量标准和样品时本底计数率的代表值。测量数据见表 F.7,其中:

t_s, t_B, t_x 分别是对标准溶液、纯水和样品溶液进行测量时,由参考时间 $t=0$ 到经死时间修正的计数时间间隔 $T_0=60$ min 的中点的时间;虽然为了完整性给出了 t_B ,但在分析中不需要它。

C_s, C_B, C_x 分别是标准溶液、纯水和样品溶液在经过死时间修正的计数时间间隔 $T_0=60$ min 内记录的计数。

观测到的计数可以表示为:

$$C_s = C_B + \epsilon A_s T_0 m_s e^{-\lambda t_s} \quad \dots\dots\dots (F.18a)$$

$$C_x = C_B + \epsilon A_x T_0 m_x e^{-\lambda t_x} \quad \dots\dots\dots (F.18b)$$

式中:

- ϵ ——对某个给定源成分的 ^{222}Rn 液体闪烁检测效率,假设其与活度等级无关;
- A_s ——在参考时间 $t=0$ 时标准的活度浓度;
- A_x ——被测量,定义为参考时间 $t=0$ 时样品的未知活度浓度;
- m_s ——标准溶液的质量;
- m_x ——被测样品的质量;
- λ —— ^{222}Rn 的衰变常数: $\lambda = (\ln 2)/T_{1/2} = 1.25\ 894 \times 10^{-4} \text{ min}^{-1}$ ($T_{1/2} = 5\ 505.8 \text{ min}$)。

表 F.7 确定未知样品活度浓度的计数数据

循环 <i>k</i>	标准		净水		样品	
	<i>t_s</i> min	<i>C_s</i> 计数	<i>t_B</i> min	<i>C_B</i> 计数	<i>t_x</i> min	<i>C_x</i> 计数
1	243.74	15 380	305.56	4 054	367.37	41 432
2	984.53	14 978	1 046.10	3 922	1 107.66	38 706
3	1 723.87	14 394	1 785.43	4 200	1 846.99	35 860
4	2 463.17	13 254	2 524.73	3 830	2 586.28	32 238
5	3 217.56	12 516	3 279.12	3 956	3 340.68	29 640
6	3 956.83	11 058	4 018.38	3 980	4 079.94	26 356

式(F.18a)和式(F.18b)表明,因为标准溶液和被测样品的活度指数衰变,从一个循环到另一个循环本底计数有轻微变化,所以表 F.7 中给出的 6 个 *C_s* 和 6 个 *C_x* 值都不能直接取平均,而是应考虑衰变修正和本底修正的计数(或者由计数次数除以 *T₀* = 60 min 定义的计数率)。这意味着综合式(F.18a)和式(F.18b),可得出下列以已知量表示的未知浓度表达式:

$$\begin{aligned}
 A_x &= f(A_s, m_s, m_x, C_s, C_x, C_B, t_s, t_x, \lambda) \\
 &= A_s \frac{m_s}{m_x} \frac{(C_x - C_B)e^{\lambda t_x}}{(C_s - C_B)e^{\lambda t_s}} \\
 &= A_s \frac{m_s}{m_x} \frac{C_x - C_B}{C_s - C_B} e^{\lambda(t_x - t_s)} \dots\dots\dots (F.19)
 \end{aligned}$$

其中, $(C_x - C_B)e^{\lambda t_x}$ 和 $(C_s - C_B)e^{\lambda t_s}$ 分别是被测样品和标准溶液在参考时间 *t* = 0 和在时间间隔 *T₀* = 60 min 时的本底修正计数,也可简写为:

$$A_x = f(A_s, m_s, m_x, R_s, R_x) = A_s \frac{m_s}{m_x} \frac{R_x}{R_s} \dots\dots\dots (F.20)$$

式中,本底修正和衰变修正的计数率 *R_x* 和 *R_s* 分别为:

$$R_x = [(C_x - C_B)/T_0]e^{\lambda t_x} \dots\dots\dots (F.21a)$$

$$R_s = [(C_s - C_B)/T_0]e^{\lambda t_s} \dots\dots\dots (F.21b)$$

F.4.2 数据分析

表 F.8 汇总了本底修正和衰减修正计数率 *R_x* 和 *R_s* 的值,这些值是用表 F.7 中的数据和前面给出的 $\lambda = 1.258\ 94 \times 10^{-4} \text{ min}^{-1}$,根据式(F.21a)和式(F.21b)计算得到的。应该注意到,比值 $R = R_x/R_s$ 可以很简单地由下式来计算:

$$[(C_x - C_B)/(C_s - C_B)]e^{\lambda(t_x - t_s)}$$

算术平均值 $\overline{R_s}$ 、 $\overline{R_x}$ 和 \overline{R} ,以及它们的实验标准差 $s(\overline{R_s})$ 、 $s(\overline{R_x})$ 和 $s(\overline{R})$,用通常方法[4.2中式(3)和式(5)]计算。相关系数 $r(\overline{R_x}, \overline{R_s})$ 由5.2.3中式(17)和5.2.2中式(14)计算。

因为 *R_x* 和 *R_s* 值的变化相对较小,平均值之比 $\overline{R_x}/\overline{R_s}$ 及其标准不确定度 $u(\overline{R_x}/\overline{R_s})$ 分别非常接近于表 F.8 最后一列给出的平均比值 \overline{R} 及其实验标准差 $s(\overline{R})$ [见 F.2.4 和其中式(F.10)]。然而,在计算标准不确定度 $u(\overline{R_x}/\overline{R_s})$ 时,由相关系数, $r(\overline{R_x}, \overline{R_s})$ 表示的 *R_x* 和 *R_s* 之间的相互关系要用5.2.2中式(16)考虑进去。[该式对 $\overline{R_x}/\overline{R_s}$ 的相对估计方差求出式(F.22b)的最后三项]

应该注意到, *R_x* 和 *R_s* 各自的实验标准差 $\sqrt{6} s(\overline{R_x})$ 和 $\sqrt{6} s(\overline{R_s})$ 显示,这些量的变化性比计数过程的

泊松统计所暗示的变化性大 2 到 3 倍；后者已包含于观测的计数变化中，不需单独考虑。

表 F.8 衰减修正和本底修正的计数率计算

循环 k	R_x min^{-1}	R_s min^{-1}	$t_x - t_s$ min	$R = R_x / R_s$
1	652.46	194.65	123.63	3.352 0
2	666.48	208.58	123.13	3.195 3
3	665.80	211.08	123.12	3.154 3
4	655.68	214.17	123.11	3.061 5
5	651.87	213.92	123.12	3.047 3
6	623.31	194.13	123.11	3.210 7
	$\bar{R}_x = 652.60$ $s(\bar{R}_x) = 6.42$ $s(\bar{R}_x)/\bar{R}_x = 0.98 \times 10^{-2}$	$\bar{R}_s = 206.09$ $s(\bar{R}_s) = 3.79$ $s(\bar{R}_s)/\bar{R}_s = 1.84 \times 10^{-2}$		$\bar{R} = 3.170$ $s(\bar{R}) = 0.046$ $s(\bar{R})/\bar{R} = 1.44 \times 10^{-2}$
	$\bar{R}_x/\bar{R}_s = 3.167$ $u(\bar{R}_x/\bar{R}_s) = 0.045$ $u(\bar{R}_x/\bar{R}_s)/(\bar{R}_x/\bar{R}_s) = 1.42 \times 10^{-2}$			
相关系数: $r(\bar{R}_x, \bar{R}_s) = 0.646$				

F.4.3 最终结果的计算

为了由式(F.20)获得未知的活度浓度 A_x 及其合成标准不确定度 $u_c(A_x)$ ，需要 A_s 、 m_x 和 m_s 及其标准不确定度。这些量为：

$$A_s = 0.136 8 \text{ Bq/g}$$

$$u(A_s) = 0.001 8 \text{ Bq/g}; u(A_s)/A_s = 1.32 \times 10^{-2}$$

$$m_s = 5.019 2 \text{ g}$$

$$u(m_s) = 0.005 0 \text{ g}; u(m_s)/m_s = 0.10 \times 10^{-2}$$

$$m_x = 5.057 1 \text{ g}$$

$$u(m_x) = 0.001 0 \text{ g}; u(m_x)/m_x = 0.02 \times 10^{-2}$$

其他可能的不确定度来源经评定认为可忽略：

——衰变时间的标准不确定度 $u(t_{s,k})$ 和 $u(t_{x,k})$ ；

—— ^{222}Rn 的衰变常数的标准不确定度 $u(\lambda) = 1 \times 10^{-7} \text{ min}^{-1}$ (显著量是衰变因子 $\exp[\lambda(t_x - t_s)]$ ，它由循环数 $k=4$ 和 6 时的 1.015 63 变化到循环数 $k=1$ 时的 1.015 70。这些值的标准不确定度是 $u = 1.2 \times 10^{-5}$)；

——与闪烁计数器的检测效率对所用源(标准、净水和样品)的可能的依赖性有关的不确定度；

——计数器死时间修正的不确定度和对计数效率对活度等级的依赖性的修正的不确定度。

F.4.3.1 结果：方法 1

如前所述， A_x 和 $u_c(A_x)$ 可以通过两种不同方法由式(F.20)求得。在第一种方法中，用算术平均 \bar{R}_x 和 \bar{R}_s 计算 A_x ，即：

$$A_x = A_s \frac{m_s \bar{R}_x}{m_x \bar{R}_s} = 0.430 0 \text{ Bq/g} \quad \dots\dots\dots (\text{F.22a})$$

将5.2.2中式(16)代入上式,得出合成方差 $u_c^2(A_x)$:

$$\frac{u_c^2(A_x)}{A_x^2} = \frac{u^2(A_S)}{A_S^2} + \frac{u^2(m_S)}{m_S^2} + \frac{u^2(m_x)}{m_x^2} + \frac{u^2(\overline{R_x})}{\overline{R_x}^2} + \frac{u^2(\overline{R_S})}{\overline{R_S}^2} - 2r(\overline{R_x}, \overline{R_S}) \frac{u(\overline{R_x})u(\overline{R_S})}{\overline{R_x}\overline{R_S}} \dots\dots\dots (F.22b)$$

式中,如 F.4.2中所指出,最后三项给出 $\overline{R_x}/\overline{R_S}$ 的估计相对方差 $u^2(\overline{R_x}/\overline{R_S})/(\overline{R_x}/\overline{R_S})^2$ 。与 F.2.4中的讨论一致,表 F.8 中的结果表明, \overline{R} 并不严格等于 $\overline{R_x}/\overline{R_S}$; 而且 $\overline{R_x}/\overline{R_S}$ 的标准不确定度 $u(\overline{R_x}/\overline{R_S})$ 并不严格等于 \overline{R} 的标准不确定度 $s(\overline{R})$ 。

将有关量值代入式(F.22a)和式(F.22b),得:

$$\frac{u_c(A_x)}{A_x} = 1.93 \times 10^{-2}$$

$$u_c(A_x) = 0.0083 \text{ Bq/g}$$

于是,测量结果可以表述为:

$$A_x = 0.4300 \text{ Bq/g}, \text{ 其合成标准不确定度为 } u_c = 0.0083 \text{ Bq/g}.$$

F.4.3.2 结果:方法 2

在第二种方法中,避免了 $\overline{R_x}$ 和 $\overline{R_S}$ 之间的相关,用算术平均 \overline{R} 计算 A_x 。于是:

$$A_x = A_S \frac{m_S \overline{R}}{m_x} = 0.4304 \text{ Bq/g} \dots\dots\dots (F.23a)$$

$u_c^2(A_x)$ 简单地表示为:

$$\frac{u_c^2(A_x)}{A_x^2} = \frac{u^2(A_S)}{A_S^2} + \frac{u^2(m_S)}{m_S^2} + \frac{u^2(m_x)}{m_x^2} + \frac{u^2(\overline{R})}{\overline{R}^2} \dots\dots\dots (F.23b)$$

由此可得:

$$\frac{u_c(A_x)}{A_x} = 1.95 \times 10^{-2}$$

$$u_c(A_x) = 0.0084 \text{ Bq/g}$$

于是,测量结果可以表述为

$$A_x = 0.4304 \text{ Bq/g}, \text{ 合成标准不确定度为 } u_c = 0.0084 \text{ Bq/g}.$$

u_c 的有效自由度可以按 F.1.6所述的方法用韦尔奇—萨特思韦特公式来评定。

如 F.2 中一样,这两个结果中最好选第二个,因为它避免了用两个量的平均值之比代替两个量之比的平均值而带来的近似性,而且它更好地反映了所采用的测量程序(事实上数据是在分立的循环中采集的)。

尽管如此,由这两种方法所得的 A_x 值之间的差异,与任意一个的标准不确定度相比都明显地小,而且两个标准不确定度的差异是完全可以忽略的。如此的一致性表明,当适当地计入了观测的相关性时,两种方法是等效的。

F.5 方差分析

本例提供了方差分析(Analysis of Variance,简称 ANOVA)方法的简要介绍。这些统计技术用来识别和量化测量中的各项随机影响,以便在对测量结果进行不确定度分析时适当地将这些影响考虑进去。虽然方差分析方法适用于广泛的测量领域,例如用于参考标准,如齐纳电压标准和质量标准等的校准,以及标准物质的定值,但方差分析方法本身并不能识别可能存在的系统影响。

在方差分析这个通用名称下包括许多不同的模型。由于其重要性,本例中讨论的具体的模型是平衡蜂窝设计。本模型的数值说明涉及齐纳电压标准的校准,其分析应该与各种实际测量情况相关。

方差分析方法在通过实验室间测试进行标准物质(RMs)定值时具有特殊的重要性,ISO 指南 35^[19]对此有详尽描述(这种标准物质定值的简要描述见 F.5.3.2)。由于 ISO 指南 35 中包含的大部分材料事实上具有广泛的适用性,该标准可以用来查询关于方差分析的补充细节,包括非平衡蜂窝设计。类似地也可查询参考文献[15]和[20]。

F.5.1 测量问题

考虑一台标称值为 10 V,每两周一次以稳压参考源为基准对其进行校准的齐纳电压标准,在校准期间,每 J 天对该标准的电位差 V_s 进行 K 次独立的重复观测。如果以 V_{jk} 表示在第 j 天($j=1,2,\dots,J$)对 V_s 的第 k 次($k=1,2,\dots,K$)观测结果,则此标准的电位差的最佳估计值为这 JK 次观测结果的算术平均值 \bar{V} [见4.2.1中式(3)]:

$$V_s = \frac{1}{JK} \sum_{j=1}^J \sum_{k=1}^K V_{jk} = \bar{V} \quad \dots\dots\dots (F.24a)$$

该平均值的实验标准差 $s(\bar{V})$ 是 \bar{V} 的不确定度的度量, \bar{V} 是电压标准的电位差的估计值,可由下式求得[见4.2.3中式(5)]:

$$s^2(\bar{V}) = \frac{1}{JK(JK-1)} \sum_{j=1}^J \sum_{k=1}^K (V_{jk} - \bar{V})^2 \quad \dots\dots\dots (F.24b)$$

注:在本例中,始终假设为补偿系统影响对观测值所做的所有修正的不确定度可以忽略,或者其不确定度可以在分析过程的最后予以考虑。对校准齐纳电压标准所用的稳压参考源的证书值(假设有个给定的不确定度)与工作值之间的差异的修正就是属于后一类,可以在分析过程的最后将其本身加到观测值的平均值上。因此,由这些观测值用统计方法得到的该标准的电位差估计值不一定是最终测量结果,估计值的实验标准差也不一定是最终结果的合成标准不确定度。

只有当不同日的观测值的变化量与同一日内观测值的变化量相同时,由式(F.24b)所求得平均值的实验标准差 $s(\bar{V})$ 才能恰当地评价 \bar{V} 的不确定度。如果有证据表明日间变化量显著大于所预期的日内变化量,该表达式的应用可能会导致对 \bar{V} 的不确定度的大幅度低估。于是引出两个问题:怎样确定日间的变化量(以日间方差分量表征)与日内的变化量(以日内方差分量表征)相比是否显著?如果是,应该怎样评定平均值的不确定度?

F.5.2 数值例子

F.5.2.1 表 F.9 给出了能够解决上述问题的数据,其中:

- $J = 10$ 是进行电位差观测的天数;
- $K = 5$ 是每天进行电位差观测的次数;

$$\bar{V}_j = \frac{1}{K} \sum_{k=1}^K V_{jk} \quad \dots\dots\dots (F.25a)$$

是第 j 天进行的 $K=5$ 次电位差观测值的算术平均值(有 $J=10$ 个这样的日平均值);

$$\bar{V} = \frac{1}{J} \sum_{j=1}^J \bar{V}_j = \frac{1}{JK} \sum_{j=1}^J \sum_{k=1}^K V_{jk} \quad \dots\dots\dots (F.25b)$$

是 $J=10$ 个日平均值的算术平均值,因此也是 $JK=50$ 个观测值的总平均值;

$$s^2(V_{jk}) = \frac{1}{K-1} \sum_{k=1}^K (V_{jk} - \bar{V}_j)^2 \quad \dots\dots\dots (F.25c)$$

是第 j 天进行的 $K=5$ 次观测值的实验方差(有 $J=10$ 个这样的方差估计);

$$s^2(\bar{V}_j) = \frac{1}{J-1} \sum_{j=1}^J (\bar{V}_j - \bar{V})^2 \quad \dots\dots\dots (F.25d)$$

是 $J=10$ 个日平均值的实验方差(只有一个这样的方差估计)。

F.5.2.2 观测值的日内变异性和日间变异性的一致性可以通过比较两个独立的日内方差分量(即同一

日进行的观测的方差) σ_w^2 的估计值来进行研究。

σ_w^2 的第一个估计值以 s_a^2 表示,由观测到的每天平均 \bar{V}_j 的变化求得。由于 \bar{V}_j 是 K 个观测值的平均值,在假设日间方差分量为零的情况下,其估计方差 $s^2(\bar{V}_j)$ 为 σ_w^2/K 。于是由式(F.25d)得:

$$s_a^2 = Ks^2(\bar{V}_j) = \frac{K}{J-1} \sum_{j=1}^J (\bar{V}_j - \bar{V})^2 \dots\dots\dots (F.26a)$$

这是 σ_w^2 的一个估计值,它具有的自由度为 $\nu_a = J - 1 = 9$ 。

σ_w^2 的第二个估计值以 s_b^2 表示,是由 $J = 10$ 个 $s^2(V_{jk})$ 值利用 F.3.6注中的公式求得的合并方差估计,这 10 个数值由式(F.25c)计算。因为这些数值中每一个的自由度为 $\nu_i = K - 1$,导出的 s_b^2 表达式就是它们的平均。于是:

$$s_b^2 = \overline{s^2(V_{jk})} = \frac{1}{J} \sum_{j=1}^J s^2(V_{jk}) = \frac{1}{J(K-1)} \sum_{j=1}^J \sum_{k=1}^K (V_{jk} - \bar{V}_j)^2 \dots\dots\dots (F.26b)$$

这是 σ_w^2 的一个估计值,它具有的自由度为 $\nu_b = J(K - 1) = 40$ 。

由式(F.26a)和式(F.26b)给出的 σ_w^2 的估计值分别为 $s_a^2 = (128 \mu V)^2$ 和 $s_b^2 = (85 \mu V)^2$ (见表 F.9)。由于估计值 s_a^2 是基于每天平均的变异性,而估计值 s_b^2 是基于每天观测值的变异性,它们的差异表明可能存在一种影响,其从一天到另一天是发生变化的,但当观测在任何一天内进行时,该影响是保持相对不变的。 F 检验被用于检验这种可能性,也可对日间方差分量为零的假设进行检验。

F.5.2.3 F 分布是正态分布随机变量的方差 σ^2 的两个独立的估计 $s_a^2(\nu_a)$ 和 $s_b^2(\nu_b)$ 的比值 $F(\nu_a, \nu_b) = s_a^2(\nu_a)/s_b^2(\nu_b)$ 的概率分布^[15]。参数 ν_a 和 ν_b 分别为两个估计的自由度,而且 $0 \leq F(\nu_a, \nu_b) < \infty$ 。对于不同 ν_a 和 ν_b 值以及各种 F 分布分位点, F 值被列成表。 $F(\nu_a, \nu_b) > F_{0.95}$ 或 $F(\nu_a, \nu_b) > F_{0.975}$ (临界值)的值通常理解为表明在统计意义上 $s_a^2(\nu_a)$ 大于 $s_b^2(\nu_b)$;而且,如果两个估计值的方差相同,得到像观测值一样大的 F 值的概率分别小于 0.05 或 0.025。(也可以选择其他临界值,如 $F_{0.99}$)

表 F.9 $J = 10$ 天的电压标准校准数据汇总,
每天的平均 \bar{V}_j 和实验标准差 $s(V_{jk})$ 由 $K = 5$ 次独立重复观测获得

量	第几天, j				
	1	2	3	4	5
\bar{V}_j/V	10.000 172	10.000 116	10.000 013	10.000 144	10.000 106
$s(V_{jk})/\mu V$	60	77	111	101	67
量	第几天, j				
	6	7	8	9	10
\bar{V}_j/V	10.000 031	10.000 060	10.000 125	10.000 163	10.000 041
$s(V_{jk})/\mu V$	93	80	73	88	86
$\bar{V} = 10.000\ 097\ V, s(\bar{V}_j) = 57\ \mu V$					
$s_a^2 = Ks^2(\bar{V}_j) = 5 \times (57\ \mu V)^2 = (128\ \mu V)^2, s_b^2 = \overline{s^2(V_{jk})} = (85\ \mu V)^2$					

F.5.2.4 将 F 检验应用于本数值例子,得:

$$F(\nu_a, \nu_b) = \frac{s_a^2}{s_b^2} = \frac{Ks^2(\bar{V}_j)}{s^2(V_{jk})} = \frac{5 \times (57\ \mu V)^2}{(85\ \mu V)^2} = 2.25 \dots\dots\dots (F.27)$$

其分子的自由度为 $\nu_a = J - 1 = 9$,分母的自由度为 $\nu_b = J(K - 1) = 40$ 。由于 $F_{0.95}(9, 40) = 2.12$ 和 $F_{0.975}(9, 40) = 2.45$,可以得出结论,在 5% 的显著水平上存在统计上显著的日间效应,而在 2.5% 的显著水平上没有。

F.5.2.5 如果因为 s_a^2 和 s_b^2 之间的差异不被视为在统计上是显著的,而拒绝日间效应的存在(这是一个不谨慎的决定,因为它可能导致对不确定度的低估), \bar{V} 的估计方差 $s^2(\bar{V})$ 应该由式(F.24b)计算。该种关系相当于将估计值 s_a^2 和 s_b^2 合并(即取 s_a^2 和 s_b^2 的加权平均,每一项的权分别为其自由度 ν_a 和 ν_b ,见 F.3.6 中的注)来得到观测值的方差的最佳估计。用该估计值除以观测的次数 JK ,得到观测的平均值的方差的最佳估计 $s^2(\bar{V})$ 。按该程序可得:

$$s^2(\bar{V}) = \frac{(J-1)s_a^2 + J(K-1)s_b^2}{JK(JK-1)} = \frac{9 \times (128 \mu\text{V})^2 + 40 \times (85 \mu\text{V})^2}{10 \times 5 \times (10 \times 5 - 1)} \dots\dots\dots (\text{F.28a})$$

$$s^2(\bar{V}) = (13 \mu\text{V})^2, \text{或 } s(\bar{V}) = 13 \mu\text{V} \dots\dots\dots (\text{F.28b})$$

$s(\bar{V})$ 的自由度为 $JK-1=49$ 。

如果假设对系统影响的所有修正值都已经考虑,而且所有其他不确定度分量都很小,那么校准的结果可以表述为 $V_s = \bar{V} = 10.000\ 097\ \text{V}$ (见表 F.9),其合成标准不确定度为 $s(\bar{V}) = u_c = 13\ \mu\text{V}$, u_c 的自由度为 49。

注 1: 在实际中,很可能还有其他显著的不确定度分量,应将其与由观测值统计得到的不确定度分量合成(见 F.5.1, 注)。

注 2: 通过将式(F.24b)中的双重和,以 S 表示,写为:

$$S = \sum_{j=1}^J \sum_{k=1}^K [(V_{jk} - \bar{V}_j) + (\bar{V}_j - \bar{V})]^2 = (J-1)s_a^2 + J(K-1)s_b^2$$

可以看出,关于 $s^2(\bar{V})$ 的式(F.28a)与式(F.24b)是等效的。

F.5.2.6 如果接受日间效应的存在(这是一个谨慎的决定,因为它避免了对不确定度的可能低估),并假设它是随机的,那么根据式(F.25d)由 $J=10$ 个日平均值计算得到的方差 $s^2(\bar{V}_j)$ 估计的不是如 F.5.2.2 中假设的 σ_w^2/K ,而是 $\sigma_w^2/K + \sigma_b^2$,其中 σ_b^2 是方差的日间随机分量。这意味着:

$$s^2(\bar{V}_j) = s_w^2/K + s_b^2 \dots\dots\dots (\text{F.29})$$

其中, s_w^2 是 σ_w^2 的估计, s_b^2 是 σ_b^2 的估计。由于从式(F.26b)计算得到的 $\overline{S^2(V_{jk})}$ 只取决于观测值的日内变化性,可认为 $s_w^2 = \overline{S^2(V_{jk})}$ 。因此, F.5.2.4 中用于 F 检验的比值 $\frac{Ks^2(\bar{V}_j)}{s^2(V_{jk})}$ 变为:

$$F = \frac{Ks^2(\bar{V}_j)}{s^2(V_{jk})} = \frac{s_w^2 + Ks_b^2}{s_w^2} = \frac{5 \times (57 \mu\text{V})^2}{(85 \mu\text{V})^2} = 2.25 \dots\dots\dots (\text{F.30})$$

于是,可以得出:

$$s_b^2 = \frac{Ks^2(\bar{V}_j) - \overline{S^2(V_{jk})}}{K} \dots\dots\dots (\text{F.31a})$$

$$s_b^2 = (43 \mu\text{V})^2, \text{或 } s_b = 43 \mu\text{V}$$

$$s_w^2 = \overline{S^2(V_{jk})} = (85 \mu\text{V})^2, \text{或 } s_w = 85 \mu\text{V} \dots\dots\dots (\text{F.31b})$$

因为 $s^2(\bar{V}_j)$ 恰当地反映了日内和日间随机方差分量[见式(F.29)], \bar{V} 的估计方差由 $s^2(\bar{V}_j)$ 得到,如式(F.25d)。因此

$$s^2(\bar{V}) = s^2(\bar{V}_j)/J = (57 \mu\text{V})^2/10, \text{或 } s(\bar{V}) = 18 \mu\text{V} \dots\dots\dots (\text{F.32})$$

$s(\bar{V})$ 的自由度为 $J-1=9$ 。

s_w^2 (以及 s_w) 的自由度为 $J(K-1)=40$ [见式(F.26b)]。 s_b^2 (以及 s_b) 的自由度为差值 $s_b^2 = s^2(\bar{V}_j) - \overline{S^2(V_{jk})}/K$ 的有效自由度[式(F.31a)],但对它的估计是个难题。

F.5.2.7 于是,电压标准的电位差的最佳估计值为 $V_s = \bar{V} = 10.000\ 097\ \text{V}$, $s(\bar{V}) = u_c = 18\ \mu\text{V}$,如式(F.32)所示。这个 u_c 值及其自由度 9 将与在 F.5.2.5 中 [式(F.28b)] 得到的当拒绝存在日间效应时的 $u_c = 13\ \mu\text{V}$ 及其自由度 49 进行比较。

如果可能,在实际测量中应该对明显的日间效应进行进一步的研究,以确定其产生的原因以及是否

存在使方差分析方法失效的系统影响。如在本例开始时所指出的,方差分析技术是为识别和估计由随机影响所引起的不确定度分量而设计的,它们不能提供关于由系统影响所引起的不确定度分量的信息。

F.5.3 方差分析在测量中的作用

F.5.3.1 这个电压标准的例子说明了一种通常叫做平衡一级蜂窝设计的方法。它是一级蜂窝设计,因为有一个观测值的“蜂窝”级别,这些观测有一个因子,即观测日期,在测量中发生变化。它是平衡的,因为每日进行相同次数的观测。本例中所给出的分析可以用来确定在特定的测量中是否存在“操作者影响”“仪器影响”“实验室影响”“样品影响”,甚至“方法影响”。因此,在本例中,可以想象将 J 个不同日所进行的观测用同一日由 J 个不同操作者所进行的观测来代替,于是,日间方差分量就变成与不同操作者有关的方差分量。

F.5.3.2 如 F.5 中所注,方差分析方法广泛应用于通过实验室间测试的标准物质定值中。这样的定值通常包括让一些独立的具有同等能力的实验室测量某种物质的样品以得到该物质要定值的特性。一般假设,实验室内和实验室间各个测量结果之间的差异,不论由什么原因产生,在本质上是统计的。每个实验室平均值被认为是对物质特性的无偏估计,而且通常把各实验室平均值的不加权平均值假设为该特性的最佳估计。

一次标准物质定值可能需要 I 个不同实验室参与,每个实验室对 J 个不同的该物质试样的所需特性进行测量,对试样的每次测量由 K 个独立的重复观测组成。因此,总的观测数是 $I \times J \times K$,总的样品数是 $I \times J$ 。这是一个平衡二级蜂窝设计的例子,与上述一级的电压标准例子相似。在这种情况下,有两个观测的“蜂窝”级别,这些观测有样品和实验室两个不同因子在测量中发生变化。这种设计是平衡的,因为每个样品在每个实验室中观测的次数(K)是相同的,而且每个实验室测量的样品数(J)也是相同的。与电压标准例子更进一步类似的是,在标准物质这个例子中,数据分析的目的是研究不同样品间和不同实验室间影响的可能存在,以及确定被定值的特性的最佳估计值的恰当的不确定度。与前一段一致,假设该估计值是 I 个实验室平均值的平均值,也是 $I \times J \times K$ 个观测值的平均值。

F.5.3.3 在 3.4.2 中指出了改变与测量结果有关的输入量,以使其不确定度基于观察数据进行统计评价的重要性。蜂窝设计和用方差分析方法对所得数据的分析可以成功地用于实际中所遇到的许多测量情况。

然而,如 3.4.1 中所示,由于时间和资源有限,改变所有的输入量是很难行得通的;在大多数实际测量情况下,最多也只能用方差分析方法评价几个不确定度分量。如 4.3.1 中所指出,许多分量需要利用输入量的可能的变化性的所有可得到的信息进行科学判断来评定。在许多情况下,一个不确定度分量,如由于样品间影响、实验室间影响、仪器间影响,或者操作者间影响而产生的不确定度分量,不能用对一系列观测值的统计分析来评定,而应根据可供利用的所有信息来评定。

F.6 基于参考标度的测量:硬度

硬度是以一种测量方法为参考才能被量化的物理概念,它没有独立于测量方法的单位。“硬度”这个量与常规的可测量的量不同,它不能被输入代数方程去定义其他的可测量的量(虽然有时它也被用于经验公式中说明硬度与某类材料的其他特性的关系),其大小由传统的测量——即对材料块或样块的压痕尺寸的测量来确定。这种测量是根据技术标准进行的,标准规定了“压头”、应用“压头”的设备结构以及设备的操作方法等。这些技术标准不止一个,所以硬度的标准也不止一个。

报告的硬度是测得的压痕尺寸的函数(取决于标度)。在本节给出的例子中,它是 5 次重复的压痕深度的均值或算术平均值的线性函数,但对于有些其他标度,该函数是非线性的。

复现量值的标准装置作为国家标准保存,某一测量装置和国家标准装置的比对是通过传递标准块进行的。

F.6.1 测量问题

在本例中,材料样块的硬度是按“洛氏(Rockwell)C”标度,用经国家标准装置校准过的测量装置确定的。洛氏 C 硬度标度单位是0.002 mm,在该标度下,硬度定义为 100(0.002 mm)减去以 mm 为单位的 5 次压痕深度的平均值。该量除以洛氏标度单位0.002 mm 所得的值称为“HRC 硬度指数”。本例中,该量简称为“硬度”,以符号 $h_{\text{Rockwell C}}$ 表示,以洛氏长度单位表示的硬度数值称为“硬度指数”,以 $H_{\text{Rockwell C}}$ 表示。

F.6.2 数学模型

测量硬度的设备(或称校准装置)在样块上所做压痕深度的平均值,应加上修正值,以确定用国家标准装置对同一块所形成的压痕深度的平均值。因此

$$h_{\text{Rockwell C}} = f(\bar{d}, \Delta_c, \Delta_b, \Delta_s) \\ = 100 \times (0.002 \text{ mm}) - \bar{d} - \Delta_c - \Delta_b - \Delta_s \dots\dots (F.33a)$$

$$H_{\text{Rockwell C}} = h_{\text{Rockwell C}} / (0.002 \text{ mm}) \dots\dots\dots (F.33b)$$

式中:

- \bar{d} ——用校准装置在样块上所做 5 次压痕的深度的平均值;
- Δ_c ——通过传递标准块将校准装置和国家标准装置进行比较所得的修正值,它等于用国家标准装置对该标准块进行的 5 m 次压痕的平均深度减去用校准装置对同一标准块进行的 5 n 次压痕的平均深度;
- Δ_b ——传递标准块上分别用两台设备来做压痕的两个部位间的硬度差(表示为平均压痕深度之差),假设为零;
- Δ_s ——由于国家标准装置缺乏重复性和硬度量定义不完整而引起的误差。虽然 Δ_s 通常假定为零,但它具有标准不确定度 $u(\Delta_s)$ 。

由于式(F.33a)的函数的偏导数 $\partial f / \partial \bar{d}$ 、 $\partial f / \partial \Delta_c$ 、 $\partial f / \partial \Delta_b$ 和 $\partial f / \partial \Delta_s$ 都等于 -1,用校准装置测量样块所得硬度的合成标准不确定度 $u_c(h)$ 可简单地表示为:

$$u_c^2(h) = u^2(\bar{d}) + u^2(\Delta_c) + u^2(\Delta_b) + u^2(\Delta_s) \dots\dots\dots (F.34)$$

其中,为简化标记, $h = h_{\text{Rockwell C}}$ 。

F.6.3 有贡献的方差

F.6.3.1 样块的平均压痕深度 \bar{d} 的不确定度, $u(\bar{d})$

重复观测的不确定度:因为新的压痕不可能做在前面的压痕的位置上,观测值不可能严格重复。由于每一个压痕肯定是做在不同位置上,结果的任何变化包括不同位置间硬度变化的影响。因此,用校准装置在样块上的 5 次压痕平均深度的标准不确定度 $u(\bar{d})$ 取作 $s_p(d_k) / \sqrt{5}$,其中 $s_p(d_k)$ 是在已知具有很均匀硬度的样块上“重复”测量确定的压痕深度的合并实验标准差(见4.2.4)。

指示的不确定度:虽然由于校准装置的显示而对 \bar{d} 所做的修正为零,但是在 \bar{d} 中有一个由于显示分辨率(导致的深度指示不确定度而引起的不确定度,表示为 $u^2(\delta) = \delta^2 / 12$ (见 D.2.2.1)。于是, \bar{d} 的估计方差为:

$$u^2(\bar{d}) = s^2(d_k) / 5 + \delta^2 / 12 \dots\dots\dots (F.35)$$

F.6.3.2 对两台设备之差进行修正的不确定度, $u(\Delta_c)$

如 F.6.2所示, Δ_c 是对国家标准装置与校准装置之差的修正值,该修正值可以表示为 $\Delta_c = z'_s - z'$,

其中 $z'_s = (\sum_{i=1}^m \bar{z}_{s,i})/m$ 是国家标准装置对传递标准块进行的 5 m 次压痕的平均深度； $z' = (\sum_{i=1}^n \bar{z}_i)/n$ 是校准装置对同一标准块进行的 5 n 次压痕的平均深度。因此，为了比较，假设每台装置由显示分辨力所引起的不确定度可以忽略，则 Δ_c 的估计方差为：

$$u^2(\Delta_c) = \frac{s_{av}^2(\bar{z}_s)}{m} + \frac{s_{av}^2(\bar{z})}{n} \dots\dots\dots (F.36)$$

式中：

$s_{av}^2(\bar{z}_s) = [\sum_{i=1}^m s^2(\bar{z}_{s,i})]/m$ 是用标准装置所进行的 m 列压痕 $z_{s,ik}$ 中每列的均值的实验方差的平均值；

$s_{av}^2(\bar{z}) = [\sum_{i=1}^n s^2(\bar{z}_i)]/n$ 是用校准装置所进行的 n 列压痕 z_{ik} 中每列的均值的实验方差的平均值。

注：方差 $s_{av}^2(\bar{z}_s)$ 和 $s_{av}^2(\bar{z})$ 是合并方差估计值，见 F.5.2.2 中式 (F.26b) 的讨论。

F.6.3.3 对传递标准块硬度变化进行修正的不确定度, $u(\Delta_b)$

OIML(国际法制计量组织)国际建议 R 12《洛氏 C 硬度标准块的检定和校准》(已废止,本例仅供参考),要求对传递标准块的 5 次测量所得的最大和最小压痕深度的差别不应超过平均压痕深度的 $x\%$, x 是硬度等级的函数。因此,设整个标准块上压痕深度的最大差为 xz' , z' 如 F.6.3.2 中所定义,其 $n=5$ 。并设该最大差由关于平均值 $xz'/2$ 的三角概率分布所描述(以这样的假设,中心值附近的值比两端的值有更大的概率,见 4.3.9)。那么,如果在 4.3.9 的式 (9b) 中 $a=xz'/2$,则由于用标准装置和校准装置分别测量的硬度的差而对平均压痕深度进行的修正的估计方差为：

$$u^2(\Delta_b) = (xz')^2/24 \dots\dots\dots (F.37)$$

如 F.6.2 中所指出,假设修正值 Δ_b 的最佳估计值本身为零。

F.6.3.4 国家标准装置和硬度定义的不确定度, $u(\Delta_s)$

国家标准装置的不确定度和由于硬度量的定义不完整引入的不确定度一起报告为估计标准差 $u(\Delta_s)$ (一个长度量)。

F.6.4 合成标准不确定度, $u_c(h)$

将 F.6.3.1 到 F.6.3.4 中的各项不确定度代入式 (F.34), 得到硬度测量的估计方差：

$$u_c^2(h) = \frac{s^2(d_k)}{5} + \frac{\delta^2}{12} + \frac{s_{av}^2(\bar{z}_s)}{m} + \frac{s_{av}^2(\bar{z})}{n} + \frac{(xz')^2}{24} + u^2(\Delta_s) \dots\dots\dots (F.38)$$

合成标准不确定度为 $u_c(h)$ 。

F.6.5 数值举例

本例的数据汇总于表 F.10 中。

表 F.10 基于洛氏 C 标度确定样块硬度的数据一览表

不确定度来源	值
校准装置在样块上所做 5 次压痕的平均深度 \bar{d} : 0.072 mm	36.0 洛氏标度单位
5 次压痕所显示的样块的硬度指数: $H_{\text{Rockwell C}} = h_{\text{Rockwell C}} / (0.002 \text{ mm}) = [100 \times (0.002 \text{ mm}) - 0.072 \text{ mm}] / (0.002 \text{ mm})$ (见 F.6.1)	64.0 HRC
校准装置在硬度均匀的样块上所做压痕深度的合并实验标准差 $s_p(d_k)$	0.45 洛氏标度单位

表 F.10 (续)

不确定度来源	值
校准装置的显示分辨力 δ	0.1洛氏标度单位
国家标准装置在传递标准块上所做的 m 列压痕的均值的实验方差的平均值的平方根 $s_{av}(\bar{z}_S)$	0.10洛氏标度单位, $m=6$
校准装置在传递标准块上所做 n 列压痕的均值的实验方差的平均值的平方根 $s_{av}(\bar{z})$	0.11洛氏标度单位, $n=6$
在传递标准块上压痕深度的允许变化量 x	1.5×10^{-2}
国家标准装置和硬度定义的标准不确定度 $u(\Delta_S)$	0.5洛氏标度单位

这种标度是洛氏 C(Rockwell C), 记为 HRC。洛氏标度单位为 0.002 mm, 因此在表 F.10 以及下文中, 可以理解为(举例说明)“36.0洛氏标度单位”是指 $36.0 \times (0.002 \text{ mm}) = 0.072 \text{ mm}$, 它只是一种表示数据与结果的简便方法。

如果把表 F.10 中给出的有关量值代入式(F.38), 可以得到下列两个表达式:

$$u_c^2(h) = \left[\frac{0.45^2}{5} + \frac{0.1^2}{12} + \frac{0.10^2}{6} + \frac{0.11^2}{6} + \frac{(0.015 \times 36.0)^2}{24} + 0.5^2 \right] (\text{洛氏标度单位})^2$$

$$= 0.307 (\text{洛氏标度单位})^2$$

$$u_c(h) = 0.55 \text{ 洛氏标度单位} = 0.0011 \text{ mm}$$

式中, 为了计算不确定度, 取 $z' = \bar{d} = 36.0$ 洛氏标度单位是足够的。

因此, 如果假设 $\Delta_c = 0$, 样块的硬度为 $h_{\text{Rockwell C}} = 64.0$ 洛氏标度单位或 0.128 mm, 其合成标准不确定度为 $u_c(h) = 0.55$ 洛氏标度单位或 0.0011 mm。

样块的硬度指数为 $h_{\text{Rockwell C}} / (0.002 \text{ mm}) = (0.1280 \text{ mm}) / (0.002 \text{ mm})$, 或

$H_{\text{Rockwell C}} = 64.0 \text{ HRC}$, 其合成标准不确定度为 $u_c = 0.55 \text{ HRC}$ 。

除了由国家标准装置和硬度定义引起的不确定度分量 $u(\Delta_S) = 0.5$ 洛氏标度单位外, 不确定度的显著分量还有装置的重复性引起的不确定度分量 $s_p(d_k) / \sqrt{5} = 0.20$ 洛氏标度单位和传递标准块的硬度变化引起的不确定度分量 $(xz') / 24 = 0.11$ 洛氏标度单位。 u_c 的有效自由度可用韦尔奇-萨特思韦特公式按 F.1.6 中所示的方法进行评定。

附 录 G
(规范性附录)
基本符号汇编

本附录给出了本标准中所用的基本符号及其含义的说明,见表 G.1。

表 G.1 基本符号一览表

符号	说 明
a	输入量 X_i 的可能值的矩形分布的半宽度; $a = (a_+ - a_-)/2$
a_+	输入量 X_i 的上限
a_-	输入量 X_i 的下限
b_+	输入量 X_i 与其估计值 x_i 的偏差的上限; $b_+ = a_+ - x_i$
b_-	输入量 X_i 与其估计值 x_i 的偏差的下限; $b_- = x_i - a_-$
c_i	偏导数或灵敏系数; $c_i = \partial f / \partial x_i$
f	被测量 Y 和与其相关的输入量 X_i 之间的函数关系,以及输出量的估计值 y 和与其相关的输入量的估计值 x_i 之间的函数关系
$\partial f / \partial x_i$	被测量 Y 和与其相关的输入量 X_i 之间的函数对于输入量 X_i 的偏导数,对 X_i 用估计值 x_i 评定: $\partial f / \partial x_i = \partial f / \partial x_i _{x_1, x_2, \dots, x_N}$
k	包含因子,用于由输出量 y 的合成标准不确定度 $u_c(y)$ 计算其扩展不确定度 $U = k u_c(y)$,其中 U 定义为一个高包含概率的区间, $Y = y \pm U$
k_p	用于由合成标准不确定度 $u_c(y)$ 计算输出量的估计值 y 的扩展不确定度 $U_p = k_p u_c(y)$ 的包含因子,其中 U_p 定义为一个具有高的规定包含概率为 p 的区间 $Y = y \pm U_p$
n	重复观测的次数
N	与被测量 Y 有关的输入量 X_i 的个数
p	包含概率或置信水平; $0 \leq p \leq 1$
q	用概率分布描述的随机变量
\bar{q}	对随机变量 q 的 n 次独立重复观测的均值或算术平均值; q 的概率分布的期望或均值 μ_q 的估计值
q_k	随机变量 q 的第 k 次独立重复观测值
$r(x_i, x_j)$	与输入量 X_i 和 X_j 的估计值 x_i 与 x_j 有关的估计的相关系数, $r(x_i, x_j) = u(x_i, x_j) / [u(x_i)u(x_j)]$
$r(\bar{X}_i, \bar{Y}_j)$	由对 X_i 和 X_j 的 n 对独立重复观测值 $X_{i,k}$ 和 $X_{j,k}$ 确定的输入平均值 \bar{X}_i 和 \bar{X}_j 的估计的相关系数, $r(\bar{X}_i, \bar{Y}_j) = s(\bar{X}_i, \bar{X}_j) / [s(\bar{X}_i)s(\bar{X}_j)]$
$r(y_i, y_j)$	当在同一次测量中需要确定两个或两个以上被测量或输出量时,与输出量的估计值 y_i 和 y_j 有关的估计的相关系数
s_p^2	合成方差或合并方差的估计值
s_p	合并实验标准差,等于 s_p^2 的正平方根

表 G.1 (续)

符号	说 明
$s^2(\bar{q})$	平均值 \bar{q} 的实验方差; \bar{q} 的方差 σ^2/n 的估计值; $s^2(\bar{q})=s^2(q_k)/n$; 由 A 类评定得到的估计方差
$s(\bar{q})$	平均值 \bar{q} 的实验标准差, 等于 $s^2(\bar{q})$ 的正平方根; $s(\bar{q})$ 是 $\sigma(\bar{q})$ 的有偏估计值; 由 A 类评定得到的标准不确定度
$s^2(q_k)$	q 的 n 次独立重复观测值 q_k 确定的实验方差; q 的概率分布的方差 σ^2 的估计值
$s(q_k)$	实验标准差, 等于 $s^2(q_k)$ 的正平方根; q 的概率分布的标准差 σ 的有偏估计值
$s^2(\bar{X}_i)$	由对 X_i 的 n 次独立重复观测值 $X_{i,k}$ 确定的输入平均值 \bar{X}_i 的实验方差; 由 A 类评定得到的估计方差
$s(\bar{X}_i)$	输入平均值 \bar{X}_i 的实验标准差, 等于 $s^2(\bar{X}_i)$ 的正平方根; 由 A 类评定得到的标准不确定度
$s(\bar{q}, \bar{r})$	平均值 \bar{q} 和 \bar{r} 的协方差的估计值, 是对 q 和 r 进行 n 对独立重复同时观测的 q_k 和 r_k 确定的两个随机变量 q 和 r 的期望 μ_q 和 μ_r 的估计值; 由 A 类评定得到的估计协方差
$s(\bar{X}_i, \bar{X}_j)$	由 \bar{X}_i 和 \bar{X}_j 的 n 次独立重复同时观测对 $X_{i,k}$ 和 $X_{j,k}$ 确定的输入平均值 \bar{X}_i 和 \bar{X}_j 的协方差的估计值; 由 A 类评定得到的估计协方差
$t_p(\nu)$	对应给定概率 p 和自由度 ν 时由 t 分布得到的 t 因子
$t_p(\nu_{\text{eff}})$	对应给定概率 p 和有效自由度 ν_{eff} 时由 t 分布得到的 t 因子; 用于计算扩展不确定度 U_p
$u^2(x_i)$	输入量 X_i 的估计值 x_i 的估计方差。 注: 当 x_i 由 n 次独立重复观测的均值或算术平均值确定时, $u^2(x_i)=s^2(\bar{X}_i)$ 是由 A 类评定得到的估计方差。
$u(x_i)$	输入量 X_i 的估计值 x_i 的标准不确定度, 等于 $u^2(x_i)$ 的正平方根。 注: 当 x_i 由 n 次独立重复观测的均值或算术平均值确定时, $u(x_i)=s(\bar{X}_i)$ 是由 A 类评定得到的标准不确定度
$u(x_i, x_j)$	输入量 X_i 和 X_j 的两个估计值 x_i 和 x_j 的估计协方差。 注: 当 x_i 和 x_j 是由 n 对独立重复同时观测值确定时, $u(x_i, x_j)=s(\bar{X}_i, \bar{X}_j)$ 是由 A 类评定得到的估计协方差
$u_c^2(y)$	输出量的估计值 y 的合成方差
$u_c(y)$	输出量的估计值 y 的合成标准不确定度, 等于 $u_c^2(y)$ 的正平方根
$u_{cA}(y)$	由 A 类评定获得的标准不确定度和估计协方差确定的输出量的估计值 y 的合成标准不确定度
$u_{cB}(y)$	由 B 类评定获得的标准不确定度和估计协方差确定的输出量的估计值 y 的合成标准不确定度
$u_c(y_i)$	当在同一次测量中要确定两个或更多个被测量或输出量时, 输出量的估计值 y_i 的合成标准不确定度
$u_i^2(y)$	由输入量的估计值 x_i 的估计方差 $u^2(x_i)$ 产生输出量的估计值 y 的合成方差 $u_c^2(y)$ 的分量, $u_i^2(y)=[c_i u(x_i)]^2$
$u_i(y)$	由输入量的估计值 x_i 的标准不确定度产生输出量的估计值 y 的合成标准不确定度 $u_c(y)$ 的分量, $u_i(y)= c_i u(x_i)$

表 G.1 (续)

符号	说 明
$u(y_i, y_j)$	在同一次测量中确定的输出量的估计值 y_i 与 y_j 的估计协方差
$u(x_i)/ x_i $	输入量的估计值 x_i 的相对标准不确定度
$u_c(y)/ y $	输出量的估计值 y 的相对合成标准不确定度
$[u(x_i)/x_i]^2$	输入量的估计值 x_i 的估计相对方差
$[u_c(y)/y]^2$	输出量的估计值 y 的相对合成方差
$\frac{u(x_i, y_j)}{ x_i y_j }$	输入量的估计值 x_i 和 x_j 的估计相对协方差
U	输出量的估计值 y 的扩展不确定度, 它确定了具有高的包含概率的一个区间 $Y = y \pm U$, 等于包含因子 k 乘以 y 的合成标准不确定度 $u_c(y)$, $U = k u_c(y)$
U_p	输出量的估计值 y 的扩展不确定度, 它具有高的规定的包含概率 p 的区间 $Y = y \pm U_p$, 等于包含因子 k_p 乘以 y 的合成标准不确定度 $u_c(y)$, $U_p = k_p u_c(y)$
x_i	输入量 X_i 的估计值。 注: 当 x_i 从 n 次独立重复观测的算术平均值或均值得到时, $x_i = \bar{X}_i$
X_i	与被测量 Y 有关的第 i 个输入量。 注: X_i 可能是物理量或随机变量(见4.1.1, 注1)
\bar{X}_i	输入量 X_i 的值的估计值; 等于 X_i 的 n 次独立重复观测值 $X_{i,k}$ 的均值或算术平均值
$X_{i,k}$	X_i 的第 k 个独立重复观测值
y	被测量 Y 的估计值; 测量结果; 输出量的估计值
y_i	当在同一次测量中要确定两个或更多个被测量或输出量时, 被测量 Y_i 的估计值
Y	被测量
$\frac{\Delta u(x_i)}{u(x_i)}$	输入量的估计值 x_i 的标准不确定度 $u(x_i)$ 的估计相对不确定度
μ_q	随机变量 q 的概率分布的期望或均值
ν	自由度
ν_i	输入量的估计值 x_i 的标准不确定度 $u(x_i)$ 的自由度或有效自由度
ν_{eff}	$u_c(y)$ 的有效自由度, 在计算扩展不确定度 U_p 时用于获得 $t_p(\nu_{\text{eff}})$
ν_{effA}	仅由 A 类评定获得的标准不确定度确定的合成标准不确定度的有效自由度
ν_{effB}	仅由 B 类评定获得的标准不确定度确定的合成标准不确定度的有效自由度
σ^2	随机变量 q 的概率分布的方差, 用 $s^2(q_k)$ 估计
σ	一个概率分布的标准差, 等于 σ^2 的正平方根; $s(q_k)$ 是 σ 的有偏估计
$\sigma^2(\bar{q})$	\bar{q} 的方差, 等于 σ^2/n , 用 $s^2(\bar{q}) = s^2(q_k)/n$ 估计
$\sigma(\bar{q})$	\bar{q} 的标准差, 等于 $\sigma^2(\bar{q})$ 的正平方根; $s(\bar{q})$ 是 $\sigma(\bar{q})$ 的有偏估计
$\sigma^2[s(\bar{q})]$	\bar{q} 的实验标准差 $s(\bar{q})$ 的方差
$\sigma[s(\bar{q})]$	\bar{q} 的实验标准差 $s(\bar{q})$ 的标准差, 等于 $\sigma^2[s(\bar{q})]$ 的正平方根

附 录 H
(资料性附录)

本标准与 ISO/IEC 指南 98-3:2008 相比的结构变化情况

本标准与 ISO/IEC 指南 98-3:2008 相比没有技术性差异,但在结构上有部分调整,具体章条编号对照情况见表 H.1。

表 H.1 本标准与 ISO/IEC 指南 98-3:2008 的章条编号对照情况

本标准章条编号	对应的 ISO/IEC 指南章条编号
—	前言
引言	0
2	—
—	2.1
—	2.2.1
—	2.2.2
3.1	2.2.3
—	2.2.4
3.2	2.3.1
3.3	2.3.2
3.4	2.3.3
3.5	2.3.4
3.6	2.3.5
3.7	2.3.6
—	C.1
—	C.2
B.1.1	3.1
B.1.2	3.2
B.1.3	3.3
B.1.4	3.4
B.1.4.9	—
B.2	D
—	C.1
—	C.2
B.3	C.3
附录 C	附录 E
附录 D	附录 F
附录 E	附录 G
附录 F	附录 H
附录 G	附录 J
附录 H	—
—	词汇索引

参 考 文 献

[1] CIPM(1980), BIPM Proc.-Verb. Com. Int. Poids et Mesures 48, C1-C30 (in French); BIPM (1980), Rapport BIPM-80/3, Report on the BIPM enquiry on error statements, Bur. Intl. Poids et Mesures (Sèvres, France) (in English)

[2] KAARLS, R. (1981), BIPM Proc.-Verb. Com. Int. Poids et Mesures 49, A1-A12 (in French); Giacomo, P. (1981), Metrologia 17, 73-74 (in English)

注：本标准的引言(0.7)中的 INC-1(1980)建议书来自于 BIPM 的内部报告,其英文本是该建议书的最终版本,与 BIPM 网站上权威的法文文本一致。Metrologia 17 中的英文译本对 BIPM 内部报告和本标准 0.7 中的英文译本有轻微的修改。

[3] CIPM(1981), BIPM Proc.-Verb. Com. Int. Poids et Mesures 49, 8-9, 26 (in French); Giacomo, P. (1982), Metrologia 18, 43-44 (in English)

[4] CIPM(1986), BIPM Proc.-Verb. Com. Int. Poids et Mesures 54, 14, 35 (in French); Giacomo, P. (1987), Metrologia 24, 49-50 (in English)

[5] ISO 5725:1986 Precision of test methods—Determination of repeatability and reproducibility for a standard test method by inter-laboratory tests, International Organization for Standardization (Geneva, Switzerland)

注：该标准已被修订,新版本有一个新的标题:《测量方法和结果的准确度(真实性和精密度)》,并且包含 6 个部分:

第 1 部分:基本原则与定义

第 2 部分:确定标准测量方法的重复性与再现性基本方法

第 3 部分:标准测量方法精密度的中间度量

第 4 部分:确定标准测量方法正确度的基本方法

第 5 部分:确定标准测量方法精密度的可替代方法

第 6 部分:准确度值

[6] ISO/IEC GUIDE 99 International vocabulary of basic and general terms in metrology, second edition, 1993 (缩写为 VIM 或 VIM:1993)

注 1: VIM 第 3 版(VIM:2007)于 2007 年出版。

注 2: VIM 第 2 版以 ISO/TAG4、BIPM、IEC、IFCC、IUPAC、IUPAP、OIML 的名义发布,第 3 版增加了国际实验室认可合作组织(ILAC)。

注 3: VIM 第 1 版于 1984 年以 BIPM、IEC、ISO 和 OIML 的名义发布。

[7] ISO 3534-1:1993 Statistics—Vocabulary and symbols—Part 1: Probability and general statistical terms, International Organization for Standardization (Geneva, Switzerland)

注: ISO 3534-1:2006(GB/T 3358.1—2009, IDT) 替代 ISO 3534-1:1993。应注意一些术语和定义被修改了。更多信息参见该标准的最新版。

[8] FULLER, W.A. (1987), Measurement error models, John Wiley (New York, N.Y.)

[9] ALLAN, D.W. (1987), IEEE Trans. Instrum. Meas. IM-36, 646-654

[10] DIETRICH, C.F. (1991), Uncertainty, calibration and probability, second edition, Adam-Hilger (Bristol)

[11] MÜLLER, J.W. (1979), Nucl. Instrum. Meth. 163, 241-251

[12] MÜLLER, J.W. (1984), in Precision measurement and fundamental constants II, Taylor, B. N., and Phillips, W. D., eds., Natl. Bur. Stand. (U.S.) Spec. Publ. 617, US GPO (Washington, D.C.), 375-381

[13] JEFFREYS, H. (1983), Theory of probability, third edition, Oxford University Press (Ox-

ford)

[14] PRESS, S.J. (1989), Bayesian statistics: principles, models, and applications, John Wiley (New York, N.Y.)

[15] BOX, G.E.P., HUNTER, W.G., and HUNTER, J.S. (1978), Statistics for experimenters, John Wiley (New York, N.Y.)

[16] WELCH, B.L. (1936), J. R. Stat. Soc. Suppl. 3, 29-48; (1938), Biometrika 29, 350-362; (1947), *ibid.* 34, 28-35

[17] FAIRFIELD-SMITH, H. (1936), J. Counc. Sci. Indust. Res. (Australia) 9(3), 211

[18] SATTERTHWAITE, F.E. (1941), Psychometrika 6, 309-316; (1946) Biometrics Bull. 2 (6), 110-114

[19] ISO Guide 35:1989 Certification of reference materials—General and statistical principles, second edition, International Organization for Standardization (Geneva, Switzerland)

注: ISO 指南 35:2006 替代 ISO 指南 35:1989。更多信息参见该标准的最新版。

[20] BARKER, T.B. (1985), Quality by experimental design, Marcel Dekker (New York, N.Y.)

[21] JJF 1001—2011 通用计量术语及定义

[22] JJF 1059.1—2012 测量不确定度评定与表示

