

王斗文 邓云 编著

检验检测实验室质量控制技术 **top-down** 不确定度评定



中国质检出版社
中国标准出版社

检验检测实验室质量控制技术

top-down 不确定度评定

王斗文 邓云 编著

中国质检出版社
中国标准出版社
北京

图书在版编目(CIP)数据

检验检测实验室质量控制技术 top-down 不确定度评定/王斗文,邓云
编著. —北京:中国质检出版社,2017. 1

ISBN 978-7-5026-4383-6

I. 检… II. ①王… III. ①实验室管理-质量管理 IV. ①N33

中国版本图书馆 CIP 数据核字(2016)第 280431 号

中国质检出版社 出版发行
中国标准出版社

北京市朝阳区和平里西街甲 2 号(100029)

北京市西城区三里河北街 16 号(100045)

网址:www.spc.net.cn

总编室:(010)68533533 发行中心:(010)51780238

读者服务部:(010)68523946

中国标准出版社秦皇岛印刷厂印刷

各地新华书店经销

*

开本 787×1092 1/16 印张 9.75 字数 249 千字

2017 年 1 月第一版 2017 年 1 月第一次印刷

*

定价:35.00 元

如有印装差错 由本社发行中心调换

版权专有 侵权必究

举报电话:(010)68510107

前 言

检验检测是国家质量技术基础的重要组成部分,是国家治理体系的重要工具和技术支撑,也是国家鼓励发展的现代服务业、高技术服务业。

检验检测实验室为客户提供数据结果时,必须面对如何做到更加科学准确、如何避免检测风险等问题,引入不确定度概念后,对这些问题的认识得到了深化。

经过多年研究和实践,我们认为采用 top-down 不确定度评定方法,即在期间精密度测量条件下,利用质量控制技术和方法确认数据,分别将偏倚和不精密度两者带来的累积效应予以合并,即可视为不确定度的估计量。这种评定方法,具有简便与操作性强等特点,最大限度避免了由于变量间存在的协方差而导致交互效应无法估计、影响分量查找的不确定性和非正态繁琐的数学运算等问题。

本书可以为检验检测实验室管理人员、技术人员及广大科技工作者提供参考,也可以为研究不确定度评定的学者提供有益的帮助。

由于 top-down 不确定度评定方法在国内各检验检测实验室的普及运用尚未全面开始,限于我们对 top-down 不确定度评定研究水平,书中难免存在错误与疏漏之处,欢迎广大读者提出意见建议。

本书的编写和出版参考了国内外相关领域专家、学者的部分文章。

本书的编写和出版得到了德宏州质量技术监督综合检测中心、德宏州检测技术协会和中国质检出版社、中国标准出版社的支持和帮助,在此我们表示衷心的感谢!

编著者

2016 年 10 月于德宏芒市

目 录

第1章 基本概念与模型分析	(1)
1.1 不确定度的提出	(1)
1.2 通用模型	(2)
1.3 两种评定方法	(4)
1.4 不同模型的操作与处理	(6)
1.5 评定策略	(8)
第2章 有关内容的讨论	(10)
2.1 统计受控与随机误差	(10)
2.2 期间精密度测量	(11)
2.3 自相关趋势与失控准则	(12)
2.4 系统误差和偏倚	(14)
2.5 过程能力与过程性能	(15)
2.5.1 过程能力	(15)
2.5.2 过程性能	(17)
2.5.3 检测性能指标	(18)
2.6 离群值检验	(18)
2.7 小样本测量	(21)
2.8 决策错误与风险预测控制	(23)
2.8.1 基于 α 、 Δ 和 $1-\beta$ 的选择	(23)
2.8.2 接受限的 β 水平控制	(25)
2.8.3 实际差异控制与功效曲线	(26)
第3章 控制图技术的研讨	(28)
3.1 简介	(28)
3.2 AD 统计技术的研究	(31)

3.2.1	问题的提出	(31)
3.2.2	AD 统计技术	(31)
3.3	控制限确定与偏倚受控	(33)
3.3.1	数据的预处理	(33)
3.3.2	控制限的确定与趋势线叠加	(33)
3.3.3	控制限的统计合并	(34)
3.3.4	偏倚受控(标准样品、参比方法、空间研究的确认)	(34)
3.3.4.1	标准样品的确认	(34)
3.3.4.2	参比方法的确认	(34)
3.3.4.3	空间研究的确认	(35)
3.4	案例分析(汽油研究法辛烷值的测定)	(35)
第4章	多水平的评定研究	(44)
4.1	线性回归的概念解释	(44)
4.1.1	常数模型	(48)
4.1.2	比例模型	(49)
4.1.3	案例分析(食品黄曲霉毒素的测定)	(50)
4.2	室间与室内研究	(56)
4.2.1	简介	(56)
4.2.2	h/k 统计的基本要求和计算	(57)
4.2.3	案例分析(石油产品硫含量的测定)	(58)
4.3	偏倚修正函数	(61)
4.3.1	背景简介	(61)
4.3.2	CSS 检验与讨论	(62)
4.3.2.1	TSS 的计算	(62)
4.3.2.2	CSS 的计算	(62)
4.3.2.3	CSS 关系式的选择	(64)
4.3.2.4	所选 CSS 的正态性检验	(64)
4.3.2.5	不确定度评定	(64)
4.3.3	案例分析(环境空气质量细颗粒物 PM _{2.5} 的测定)	(65)
4.4	幂函数式的模型研究	(73)
4.4.1	经验模型的由来与研究意义	(73)
4.4.2	经验模型的拟合与讨论	(76)
4.4.3	案例分析(环境空气质量二氧化硫的测定)	(77)

第 5 章 检出限和定量限	(81)
5.1 背景简介	(81)
5.2 LC、IDE 和 IQE 的理解与分析讨论	(83)
5.2.1 LC、删失、 α 和 β	(83)
5.2.2 IDE 和 IQE	(85)
5.2.3 OLS 与 WLS 拟合选择	(86)
5.2.4 模型拟合的识别检验与 IDE 和 IQE 的计算	(88)
5.3 案例分析(模拟与作图)	(89)
5.3.1 IDE 评定	(89)
5.3.2 IQE 评定	(92)
第 6 章 综合应用案例	(94)
6.1 金属材料力学拉伸性能抗拉强度的测试	(94)
6.2 室内环境污染源人造板游离甲醛的测定	(97)
6.3 营养液中栽培的薄荷植株茎生长实验的室内研究	(102)
6.4 客户投诉响应的量化监控	(105)
6.5 石油产品氯含量的定量限测定	(108)
6.6 纺织服装产品定性分析量化研究	(111)
6.7 低合金钢碳含量的测定	(112)
6.8 黑色金属矿锰矿中铝含量的测定	(116)
6.9 饮用水综合指标 COD 的全年质量安全监控	(118)
6.10 FT-120 乳品分析仪法测定纯牛奶的脂肪含量	(120)
6.11 溶剂型木器涂料中甲苯含量的测定	(123)
6.12 整数规划指派问题标准形的解析	(124)
6.12.1 匈牙利法思想的效益矩阵模型	(124)
6.12.2 0 元素变动的最优解运算与模型答案	(126)
附录 A 统计数值表	(130)
表 A.1 z 数值表	(130)
表 A.2 p 数值表	(133)
表 A.3 t 分布双侧情形分位数表	(135)
表 A.4 F 分布分位数表($F_{0.95}$)	(135)
表 A.5 F 分布分位数表($F_{0.975}$)	(136)
表 A.6 95% 和 99% 概率下的 h 与 k 临界值	(136)

表 A.7 χ^2 分布单侧情形分位数表	(137)
---------------------------------	-------

附录 B 符号集注	(138)
-----------------	-------

参考文献	(143)
------------	-------

第 1 章 基本概念与模型分析

1.1 不确定度的提出

检验检测实验室测量过程中给出的精密度和正确度有助于产品质量优劣和成败的动态监控,其监控的可靠性来自于测量不确定度(不确定度)的评定。鉴于测量结果接近“真”值或接受参照值(ARV)程度是准确度,要想获取实际值,则要求给出不确定度。本书提出的 top-down 做法,即将偏倚和精密度两者的效应合并为一个估计值,其中的期间精密度标准差($s_{R'}$)可作为实验室“受控”状态能力的变量估计。

通常的受控研究属于每个实验室自己考虑的技术研究内容。GB/T 28043—2011《利用实验室间比对进行能力验证的统计方法》和 CNAS-GL02—2006《能力验证结果的统计处理和能力评价指南》作为实验室的质量评估依据。为了能合理获得方法的不确定度性能技术指标,实验室必须确保其处于统计过程控制(SPC),跟踪监控和分析系统的稳定性、精密度和偏倚是否处于动态受控状态、以及这种受控状态是否代表日常系统的性能,否则需查找系统今后要改进的具体措施,这样才会完成 GB/T 27025—2008《检测和校准实验室能力的通用要求》提出的技术要求,也是 top-down 评定的核心理念。

GB/T 27025 明确规定实验室需给出不确定度,这样的测量结果才会呈完整性和有效性。我国目前一直在推行 JJF 1059.1—2012《测量不确定度评定与表示》国家计量技术规范,该规范是参考 ISO/IEC GUIDE 98-3(2008)《Uncertainty of measurement data—Part 3: Guide to the expression of uncertainty in measurement》而制定的。实际上,早在 20 世纪 50~60 年代,不确定度的研究就已展开。

(1) 20 世纪 60~70 年代,作为计量领域的代表,NBS(美国计量科学研究院,NIST 的前身)就提出了不确定度的概念,并与美国 ASTM 给出了方法重复性、再现性和偏倚的要求,建议选择其中之一的解释:系统误差和随机误差均忽略;保留系统误差忽略随机误差;系统误差和随机误差均不能忽略;保留随机误差忽略系统误差。

(2) 化学领域将方法性能分为准确性和精密度,认为前者应尽可能予以排除,后者可通过重复来减少效应。

(3) 医学领域认为,小于限量的总误差应包括不准确性和不精密度。

(4) 1994 年, Edition Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Techniccal Note 1297、ISO GUM(ISO/IEC GUIDE 98-3 的前期版)和 ISO/IEC 17025—2005 General requirements for the competence of testing and calibration laborato-

ries 出台,规定实验室要评定和报告不确定度。

(5)1979 年,从计量、校准和物理的角度上,国际上提出了量值溯源的不确定度需求。

那么,究竟怎样来评定检测实验室的不确定度呢?国际和国内的有关文件提到了如下观点和看法。

(1)ISO JCGM 104(2009) Evaluation of measurement data—An introduction to the“Guide to the expression of uncertainty in measurement and related documents”指出:“不确定度属于质量保证(QA)的核心概念”,所以,不确定度的评定应该与质量控制(QC)活动紧密结合起来。

(2)ISO/IEC GUIDE 98-3 指出:“GUM 规定了通用原则而非技术程序,有必要出台各自领域下的评定标准”。

(3)ISO TC69 指出:“为避免不确定度低估的风险,可以根据长期的变异来对输出结果直接进行监控”。

(4)CNAS-GL06—2006《化学分析中不确定度的评估指南》指出:“实验室引入不确定度评估程序时,应该与其现有的 QA 措施结合起来,因为 QA 措施通常提供了评估不确定度所需要的很多信息”。

(5)GB/T 27411—2012《检测实验室中常用不确定度评定方法与表示》指出:“JJF 1059 正在广泛应用于各类检测实验室;在应用过程中,实验室尤为关注的是操作性强、实用而便捷的评定方法。本标准是在满足特定条件下,对 JJF 1059 的简化和延伸应用”。

上述国内外的观点无外乎均传递了这样一个理念,即在质量管理(QM)活动中,不确定度评定应该与 QA 和 QC 紧密结合起来。实际上,QA 和 QC 的某些活动是相互关联的,两者之间无本质上的区别,最终都是向客户负责,只是从管理和技术上的认识层面有所不同。所以,无论是提供外部的 QA 以及实施内部的 QM,都要从 QC 角度上着手。本书至始至终都在强调 QC 活动前提下的不确定度研究,至于 QM、QA 和 QC 三者之间的微妙区别,本书不予做过多的描述。

top-down 评定的统计动态跟踪监控的理念始终贯穿于 QA 和 QC 活动中,从总体意义上看,后者可称其为实验室的规范操作,属于实验室的一个关键组织基础,这种基础对分析测量系统的可靠性格外予以注意。为了能达到这种水准要求,实验室制定了许多相应的活动内容,而且不断地对其进行修正和更新,并形成系统化文件。然而,这决不意味着实验室由此就可以保证获得满意的质量数据来。这样就面临着一项很重要的工作内容,即检测实验室统计过程控制 SPC 活动的开展。应该认识到,不确定度的评定来自于 SPC,但确保实验室的 SPC 并非易事。从理论上讲,若参加实验室能给出优良的能力验证(PT)质量数据,则该实验室的 QC 实施得就好,给出的不确定度就会合理,而且在今后 PT 中同样会有潜力给出其他相应的质量数据;反之亦然。

1.2 通用模型

理想情况下的测量响应仅随被测量而定,但这种响应还要考虑到测量系统、样品和标准的

影响。测量系统涉及程序和试剂、仪器设备和人员操作等特性。被测样品和标准涉及变异和校准效应,由此会产生以下不同类型的误差。

(1) 由于操作者、仪器、实验室、环境等因素,使得影响因素的平均值通常会在测量子体之间变动;如果对每个子体都进行校准,因子体内的变异响应,所实施的校准会引进随机校准误差。若真实的关系呈曲线时却使用线性模型,则校准曲线的模型误差会出现系统效应,见 GB/T 22554—2010《基于标准样品的线性校准》。

(2) 未知样品的影响量不同于标准时,测量系统的性能也会不同,即由于被测量和响应之间的关系不同而产生系统误差。假设影响量之间再出现交互作用,那么子体间的变异单靠校准是不会消除的,可参见 GB/T 27408—2010《实验室质量控制 非标准测试方法的有效性评价 线性关系》。如果样品间的影响量不同,则每个样品就有可能产生系统误差,这种误差仅靠重复测量是不会减少的。而且针对不同样品的变异,影响量的交互作用同样会导致不同差异。

(3) 如果影响量可测且其结果效应已知,则有可能进行修正,较小的修正误差即可替代原误差。实际上,影响量无法得知也不可能观察到,且经校准后的修正还具有以下分量:

- 同一校准曲线使用下子体或批次内的随机变异;
- 因校准和重复测量而产生的批次间随机变异;
- 由操作者、仪器、试剂批等所导致的批次间变异;
- 室内的变异导致室间的波动,由此引起室间的平均水平波动;
- 样品变异依赖于水平被测量的函数关系;
- 每个样品出现的系统或随机效应;
- 诸多变异随时间渐进而非呈独立性变化。

如果上述误差分量呈独立性,则样品测量结果(y)的标准不确定度,即 $u(y)$ 按式(1-1)表述:

$$u(y) = \sqrt{u^2(\beta) + u^2(\delta) + u^2(d) + u^2(b_{\text{lab}}) + u^2(b_{\text{time}}) + u^2(b_{\text{run}}) + u^2(\varepsilon)} \quad (1-1)$$

式中: $u(y)$ —— y 的标准不确定度;

$u(\beta)$ ——系统误差修正后的分量;

$u(\delta)$ ——偏倚分量;

$u(d)$ ——样品批次的随机误差分量;

$u(b_{\text{lab}})$ ——实验室的系统误差分量;

$u(b_{\text{time}})$ ——时间变异的系统误差分量;

$u(b_{\text{run}})$ ——批次的校准系统误差分量;

$u(\varepsilon)$ ——批次的随机误差分量。

为了获取不确定度区间(包含区间),标准不确定度要乘上一个因子 k ,给出扩展不确定度 $U(y) = k \cdot u(y)$;其中, k 值的选择通常为 2~3。

测量结果的误差结构属于加法或线性模型。当误差分量的标准差依赖于被测水平时,即

相对标准差独立于被测值,此时采用乘法模型更合适,即误差结构是以相对而非绝对来表述。有时,结果的对数变换可将乘法误差模型变成加法。如果批次来自于重复测量条件但没有实施校准环节时,也可使用式(1-1)模型。

实际上绝大部分检测方法是基于室间研究来完成的,这种研究给出了方法的重复性、复现性和正确度的估计,由此提供了室内和室间的方差分量,也提供了正确度的不确定度估计,详见 GB/Z 22553—2010《利用重复性、再现性和正确度的估计值评估测量不确定度的指南》。室间研究假定,所识别的、不可忽略的系统效应已被修正或排除,对于未涵盖的所有因素都应以识别,由此产生的相关方差都要量化。所以,检测实验室可通过核查自身的偏倚和精密度,以确保所使用的方法符合其规定的性能。由此,如果偏倚和精密度处于受控的前提条件下,则更为简捷且有效的通用模型可由式(1-2)给出:

$$u^2(\gamma) = u^2(\delta) + s_{R'}^2 \quad (1-2)$$

式中: $u(\delta)$ ——偏倚的不确定度, $u(\delta) = \sqrt{s(\delta)^2 + u(\mu)^2}$,其中, $s(\delta)$ 为平均值的不确定度,通过计算式

$$s(\delta) = \sqrt{\left(\frac{s_{R'}^2 - \left(\frac{n-1}{n} \right) s_r^2}{N} \right)^2}$$

求得, n 和 N 分别为重复数和水平数; $s_{R'}$ 为期间精密度标准差; s_r 为重复性标准差; $u(\mu)$ 为ARV的不确定度。

当 $u^2(\delta)$ 可做忽略时,可直接取 $s_{R'}$,这种评定步骤如下:

- (1)在期间精密度测量条件下,对类似基体和水平下的被测对象数据对差进行统计和监控,以减少外界条件变化所产生的不同影响;
- (2)利用 Anderson Darling (AD) 统计,对系列数据对差进行正态性和独立性检验;
- (3)被测对象可溯源至标准样品(RM)、能力验证(PT)或参比方法,并对系列差值进行 t 检验;
- (4)利用质量控制样品(QC样)求得 $s_{R'}$ 的估计值;
- (5)在确保系统受控与偏倚效应忽略的情况下, s_r 与 $s_{R'}$ 的统计合并即为不确定度估计值。

1.3 两种评定方法

评定不确定度有以下两种方法。

(1) top-down 方法

测量系统作为一个“黑匣子”,将模型中如下的各个参数分成系统误差和随机误差:

——内部质量控制(IQC)、能力验证(PT)或室间研究(ILC);

- 有证标准样品 (CRM) 的测量;
- 标准样品 (RM) 的加标回收;
- 与参比方法比对测量。

(2) bottom-up 方法

根据不同的输入量 (响应值、修正值、给定标准值等), 建立实际或假设的误差传播模型, 估计和传播这些输入量的不确定度。

图 1-1 给出了两种评定方法的图解。

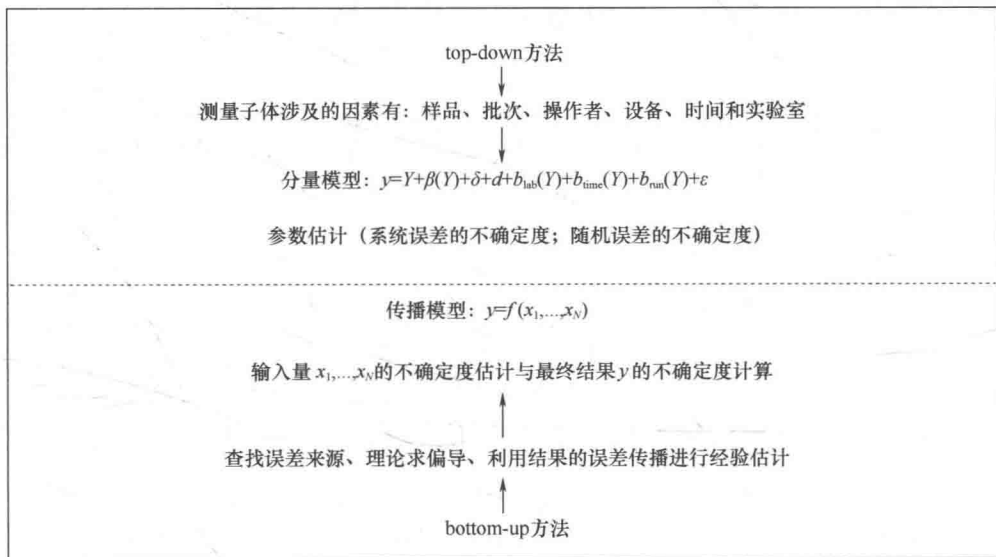


图 1-1 不确定度评定的两种方法

图中 $\beta(Y)$ 为系统误差修正值; δ 为总体效应; d 为样品批次随机效应; $b_{\text{lab}}(Y)$ 为实验室效应;

$b_{\text{time}}(Y)$ 为室内变异时间分量; $b_{\text{run}}(Y)$ 为批次内分量; e 为随机误差

两种评定方法都可结合使用进行互补, 但两者都具有不同的局限性:

bottom-up 方法的不足在于, 模型中仅包含已知的误差源, 由此会导致不确定度低估的风险;

top-down 方法的不足在于, 较难获得目标总体的代表性样品。

bottom-up 方法可以识别不同误差源的不确定度贡献, 有助于方法优化; top-down 方法的模型涵盖了测量总体, 有助于结果的解释和处理。

bottom-up 方法基于每个输入量对不确定度的贡献, 模型中含有诸多输入影响量。这些影响量的修正不仅暗指它们的识别, 还要考虑修正模型中的各种参数, 其中, 灵敏系数可通过理论求偏导, 但有时也只能依靠稳健测试给出, 然后贡献到不确定度的估计中。由此, 不仅要识别出影响量的不确定度, 还不能忽略各种参数的不确定度。实际上, 这项工作难以完成, 尤其是校准溶液和未知样品间的不同特性必然产生交互作用, 若样品间的特性又发生变化, 则会导致模型中各种参数同样变化。所以说, bottom-up 方法中的 $u(y)$ 物理解释不清晰, 因其包

含了靠判断的估计。

在不确定度评定过程中,国际上很多标准和文件也都在极力地推荐控制图技术的使用。例如,ISO/IEC GUIDE 98-3 就提到,要识别所有影响量并建立完整的数学模型,通常很难做到。若条件允许,可利用控制图和核查标准,确保测量处于统计受控状态,最终给出可信的不确定度评定。所以说,top-down 方法是基于服从假定分布的随机变量, $u(y)$ 更容易观测和测量到,从而给出更直接的解释。

1.4 不同模型的操作与处理

当对识别的影响量做可能的修正后,即可建立理想和相关的模型结构,前者是指足可逼近实际结构,后者基于方法的应用来做分配。见以下测量系统中四种情况的分析:

- ①批次内重复测量处于随机波动且非呈显著趋势;
- ②校准液和样品等同处理;
- ③校准液和样品处于相同条件,其响应和浓度关系具一致性;
- ④所有样品处于相同条件下的等同处理,其响应和浓度关系具一致性。

情况①可通过随机性检验来做调查。有时可建立校准曲线,并通过质量控制样品(QC样)和核查标准样品(CS样)的控制图使用,来检查响应和浓度之间关系的样品结果是否出现变化。

情况②需要确认。

较难确认是否符合③和④的情况,唯一的办法只能做以下调查:

系统的变化(仪器、性能或操作)是否对校准液和样品同步,为了涵盖未识别的影响效应,可做析因试验;样品稀释和加标后,是否能获得一致性结果。

(1) 满足①的理想模型

当上述四种情况中的所有条件均得到满足时,即视为理想情况下给出的模型。此时的标准液和样品均属于等同操作和等同处理,可采用式(1-3)模型:

$$y = Y + \beta_{\text{cal}}(Y) + \beta_{\text{curve}}(Y) + b(Y) + \varepsilon \quad (1-3)$$

式中: $\beta_{\text{cal}}(Y)$ ——来自校准液给定值的系统误差,可从证书或制备过程中获知;

$\beta_{\text{curve}}(Y)$ ——来自校准曲线拟合的失拟误差,其赋予标准不确定度为 $\frac{D_{\text{max}}}{\sqrt{3}}$,其中 D_{max} 是标准液稀释的最大偏离。

$b(Y), \varepsilon$ ——分别为系统误差和随机误差。随机响应的标准差估计,可用多批次的批内QC样重复测量,其中, ε 涉及纯物质或校准液制备,制备过程中的随机误差可纳入 $b(Y)$ 、校准过程中的系统误差可纳入 $\beta_{\text{cal}}(Y)$ 。

对于简单的单点校准模型而言,批内样品和标准液的响应关系有: $y = (p/q)m$,其中, p 为样品响应; q 为标准液响应; m 为标准液的给定浓度。 y 以输入量的显函数表述,假设响应依赖

于浓度。

$u(y)$ 可通过输入量的不确定度计算。当 y 为输入量的乘法函数时, 用相对不确定度 w 替代 u , 则有: $w(y) = \sqrt{w^2(\bar{p}) + w^2(q) + w^2(m)}$

其中, m 等同于 $\beta_{\text{cal}}(Y)$ 不确定度。

为测定 p 和 q 的不确定度, 其 m 的目标值必须确定。 p 和 q 的不确定度由批次内重复量的 w 表述, 相当于 $b(y)$ 和 ε 。至于 $\beta_{\text{curve}}(Y)$ 不确定度, 已假设结果中不存在系统误差。如果引入假设修正, 则会有: $y = a_{\text{curve}}(p/q)m$ 。由此, y 的相对不确定度可表示为: $w(y) = \sqrt{w^2(a_{\text{curve}}) + w^2(p) + w^2(q) + w^2(m)}$, 其中, a_{curve} 表示修正因子最佳估计为 1 时的不确定度。

(2) 未满足②的模型

如果被稀释未知样的测量程序中没有进行校准, 则会引进系统和随机分量。若这些效应独立于不同样品特性时, 可按如下简单的示例进行评定:

$y = h \times a_{\text{curve}}(p/q)m$, 其中, 稀释因子常数 h 对 p 有影响, p 目标值的合理确定可视为稀释响应的期望值, 其涵盖了稀释误差的效应分量, 则可对 h 赋予不确定度。

然而, 稀释误差视为系统误差(含子体的恒定系统误差和批内随机误差), 无法通过校准来消除。由此, y 的相对不确定度可表示为:

$$w(y) = \sqrt{w^2(b_h) + w^2(\varepsilon_h) + w^2(a_{\text{curve}}) + w^2(p) + w^2(q) + w^2(m)},$$

式中, b_h 和 ε_h 分别代表批次内系统误差和随机误差(不确定度信息来自稀释器具), 稀释会改变样品特性, 若使用同一器皿稀释时, 可消除系统误差效应。然而, 如果存在依赖于样品特性的效应, 则无法从理论上进行推理, 必须依靠试验做估计。

(3) 未满足③的模型

如果存在样品影响量, 或未校准而使得校准液与样品的浓度及响应间产生不同关系, 由此测量的结果会出现系统误差。系统误差修正和不确定度评定可分别采用 bottom-up 法和 top-down 法。

bottom-up 需要考虑所有关键的影响量修正, 包括实际和假设修正, 都应在误差模型中体现出来。实际上, 要识别出所有关键的影响量并建立正确的修正模型, 通常是无法做得到的。

top-down 属于对直接观测结果进行不确定度评定, 其模型可简化为式(1-4)。

$$y = Y + \beta(Y) + b_{\text{lab}}(Y) + b_{\text{time}}(Y) + b_{\text{run}}(Y) + \varepsilon \quad (1-4)$$

式中: $\beta(Y)$ ——被测量的恒定函数, 需靠修正来排除, 但作为系统误差通常不被所知, 只能用修正误差来替代。同标准差的处理方法相同, 求出其标准差或标准不确定度。

系统误差的修正和估计可考虑如下三种方案。

第一种方案:

实验室的系统误差 $\beta(Y) + b_{\text{lab}}(Y)$, 通过标准样品 RM 的长期测量, 估计出修正值的不确

定度、以及 $b_{\text{time}}(Y)$ 、 $b_{\text{run}}(Y)$ 和 ε (所有不确定度和标准差可视为 Y 的函数) 的标准差。

第二种方案:

参加方法的协同实验,同批次下每个实验室提交双实验结果,根据汇集结果求得方法的系统差 $\beta(Y)$ 估计、以及 $[b_{\text{lab}}(Y) + b_{\text{time}}(Y) + b_{\text{run}}(Y)]$ 和 ε (假设所有实验室操作相同) 的标准差。系统误差的估计可用来作结果修正,然后将其和随机分量予以合并。若实验室所用方法等同其他实验室(室内精密度相同、实验室偏倚来自同一总体),即认为其结果不确定度有效。

第三种方案:

实验室若能持续参加 PT 计划方案,就有可能根据平均水平而发现其是否出现偏倚,进而对方法偏倚 $\beta(Y)$ 和实验室偏倚 $b_{\text{lab}}(Y)$ 进行修正。然后,实验室就可以将这些修正不确定度与 $s_{R'}$ (期间精密度标准差) 或 s_R (复现性标准差) 一并予以合并。

(4) 未满足③和④的模型

某些情况下,即便无法满足④,但会满足③。可分别采用 bottom-up 法和 top-down 法进行处理。

bottom-up 不能对样品影响量变异间的交互进行修正。然而,模型中必须纳入样品的影响量和估计参数。但这些影响量的交互效应很难识别、也无法观察到,由此要想获取可信的不确定度比较难。

top-down 可以通过实验室间比对来解决,利用参比方法的期望值,即可求得所研究方法的不确定度,其合理模型见式(1-5):

$$y = Y + b_{\text{lab}}(Y) + b_{\text{time}}(Y) + b_{\text{run}}(Y) + \varepsilon \quad (1-5)$$

1.5 评定策略

图 1-2 给出了不确定度评定的策略概要。

基于理论判断或实验结果,来确定前述的四种不同情况。

若所有的判断情况具有其合理性,则可对不确定度进行评定。

如果处于理想状态下,但出现一个或多个问题时,可分别利用 bottom-up 法和 top-down 法进行解决。

然而,若测量系统中样品和交互因素之间出现影响量变异时,bottom-up 法的处理很难符合实际情况。

即便在所有测量条件下(无论是否经过校准),如果测量可以重复、且能获得标准样品或参比方法,都可采用 top-down 法予以解决。

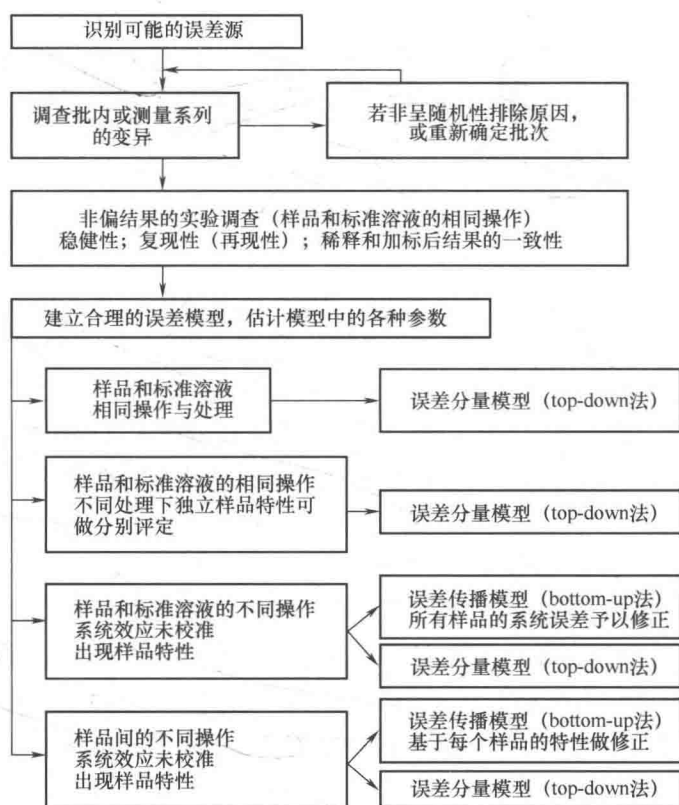


图 1-2 化学定量分析方法不确定度评定的策略概要