

钱钟泰 编著

执行

测量不确定度表示指南

ISO 1993(E)

的问题及解决方法



ISO
1993
(E)

执行“测量不确定度表示指南 ISO 1993(E)”的问题及解决方法

钱钟泰 编

中国计量出版社

图书在版编目(CIP)数据

执行“测量不确定度表示指南 ISO1993(E)”的问题及解决方法/钱钟泰编. —北京: 中国计量出版社, 1998

ISBN 7-5026-1138-X

I. 执… II. 钱… III. 不确定度-测量 IV. TB9

中国版本图书馆 CIP 数据核字(98)第 29744 号

内 容 提 要

本书由十篇论文组成, 就不同侧面对 BIPM 等七个国际组织在 1993 年制定的“测量不确定度表示指南 ISO1993(E)”中的问题提出作者的见解, 并指出解决问题的方法。形成了独具风格的与“'93 国际指南”兼容的误差评估方法的理论体系。

本书作者有自己独到的见解, 所提出的理论观点可供从事“测量不确定”的使用、研究人员和教学人员参考。

中国计量出版社出版

北京和平里西街甲 2 号

邮政编码 100013

北京市迪鑫印刷厂印刷

新华书店北京发行所发行

版权所有 不得翻印

*

787mm×1092mm 16 开本 印张 10.75 字数 261 千字

1999 年 2 月第 1 版 1999 年 2 月第 1 次印刷

*

印数 1—1000 定价: 17.50 元



前 言

本书是研究国际计量局（简称为 BIPM）等七个国际组织制定的“测量不确定度表示指南 ISO 1993(E)”在执行中的问题及其解决方法的论文集，这些论文是为了向 BIPM 等七个国际组织反映“指南”的问题而编写的。

20 年来，BIPM 等国际组织一直致力于建立准确度评定的国际等效性。1993 年同时制定了“国际通用计量学基本术语”（第二版）（下文简称“VIM'93”）及“测量不确定度表示指南 ISO 1993(E)”（下文简称“'93 国际指南”）两个文件，正是这种努力的重要一步。这种努力无疑是应该肯定的。但笔者对“'93 国际指南”的研究表明，它存在严重的概念混乱和缺陷，实际上是不可执行的。同时研究表明，并不存在弄不清楚和不可解决的问题。只要对实践中的“测量准确度的实践控制方法”研究和综合归纳，就可以制定出概念明确、切合实际、普遍适用的误差评估方法。这正是制定“'93 国际指南”的目标。这一问题的最终解决有待于 BIPM 等七个国际组织采取措施。为了向他们反映问题，编写了本书的 10 篇论文。这 10 篇文章指明和解决了“'93 国际指南”的问题，提供了一个概念明确、切合实际、消除执行障碍、规范化，且与“'93 国际指南”兼容的误差评估方法，形成了独特的理论体系。

参加本书论文编写的有下列人员：

中国计量科学研究院：钱钟泰、童光球、于渤、何 强、邹本霞；

中国计量学院：宋明顺、顾龙方；

清华大学：朱鹤年；

哈尔滨理工大学：马怀俭、王学伟；

哈尔滨电表研究所：费宇航。

1997 年 10 月国际计量局局长 QUINN 博士访问北京时，编者曾向他反映了“'93 国际指南”有关问题，并按他的要求，给他寄去前 9 篇文章提要的英文稿，这些提要英文稿由中国技术监督情报研究所康金玉、鲍建忠翻译，中国计量科学研究院童光球核校。在此谨向译者致谢。

本书各论文有关内容更详细的说明可见本书的“代序”。

编 者

1997 年 11 月

目 录

执行“测量不确定度表示指南 ISO 1993(E)”的问题及解决方法(代序)	钱钟泰 (1)
一、从术语到实例看 BIPM,IEC,IFCC,ISO,IUPAC,IUPAP 和 OIML 七个国际组织制定的“测量不确定度表示指南 ISO 1993(E)”中的概念混乱及改善其体系的建议	钱钟泰、宋明顺、顾龙方、邹本霞 (10)
二、误差的随机特性与它区别于其他随机变量的特殊性以及对“测量不确定度表示指南 ISO 1993(E)”的“测量不确定度评定”方法的评论	钱钟泰、朱鹤年、童光球、马怀俭、王学伟、何强 (19)
三、工程测量中的实用误差评估方法	钱钟泰、朱鹤年、童光球、马怀俭、王学伟、费宇航 (30)
四、BIPM,IEC,IFCC,ISO,IUPAC,IUPAP 和 OIML 七个国际组织制定的“测量不确定度表示指南 ISO 1993(E)”中有关实例的重新评估	钱钟泰、宋明顺、顾龙方、邹本霞 (49)
五、不同测量目的测量列误差评估的完整分析	钱钟泰、宋明顺、顾龙方、邹本霞 (69)
六、工程测量中与“准确度控制”措施密切结合的分层次的误差分解方法	钱钟泰、朱鹤年、童光球、马怀俭、王学伟、费宇航 (79)
七、误差极限值及其直接综合方法	钱钟泰、朱鹤年、宋明顺、顾龙方、何强 (98)
八、多次组合测量数据处理的一般方法及其合理性的证明	钱钟泰 (122)
九、与“测量误差”有关基本术语的修改意见	钱钟泰、于渤、宋明顺、顾龙方、何强 (132)
十、系统误差的非数据处理(B类)评定方法	钱钟泰、童光球、宋明顺、顾龙方、马怀俭、王学伟 (160)
参考文献	(167)

执行“测量不确定度表示指南 ISO 1993(E)”的问题及解决方法

(代 序)

钱钟泰

国际计量局(下文简称BIPM)等七个国际组织于1993年同时制定了“国际通用计量学基本术语”(第二版)(下文简称“VIM '93”)及“测量不确定度表示指南 ISO 1993(E)”(下文简称“'93国际指南”)两个文件,其内容在计量学中是属于最基础性的。“VIM '93”的内容“通用计量学基本术语”是计量信息交流的学术基础,“'93国际指南”的内容是误差评估的重要组成部分。而误差评估是计量及质量控制所有量化工作的学术基础。BIPM等七个国际组织制定这两个文件,是在国际上统一这两个重要学术内容的一种努力,这种努力无疑是值得肯定的。

但“VIM '93”和“'93国际指南”两个文件的内容是不协调的。“'93国际指南”在它的附录B中引用“VIM '93”的有关术语时,加了多达九处的“指南评注”。评注的核心是否定“VIM '93”的“[量的]真值”及“[测量]误差”的词条。“'93国际指南”还在它的以“真值、误差和不确定度”为标题的附录D中表述了下列观点:

在其附录D一开始就指出,指南的【不确定度概念是基于测量结果及其评定不确定度,而不基于不能知道的“真值”及误差】。

在其D.3.5款中的内容如下:

【D.3.5 术语“被测量真值”或量的真值(常简称真值)在本指南中避免使用,因为“真”视为多余……】

在这里,它认为被测量“真”值就是被测量值。而却在D.2款中引入了“认得量”的词条。其内容如下:

【D.2 认得量

D.2.1理想地,测量得的量应与被测量定义相一致,后者经常不能认识,而测量是对被测量的近似量进行。】

而在附录D的D.4款中误差定义如下:

【D.4 误差

由于观测的随机变化(随机效应),确定系数效应修正的不完善,一些物理现象(也是系统效应)知识的不完善,故认得量的测量是不完善的。已修正结果并不是被测量值,即存在误差。……】

“VIM '93”引用的与“不确定度”有关的词条仅“测量不确定度”一条:

【3.9 测量不确定度

与测量结果相关的参数,表征合理地赋予被测量的值的分散性

注:

1. 此参数可以是,如标准差(或其倍数),或置信区间的半宽度。

2. 测量不确定度一般由许多成分组成。一些成分可以由测量列结果统计分布估计,由实验标准差表征。另一些也可用标准差表征的成分是基于实验或其他信息的概率分布来估计。

3. 应这样理解测量结果。它是被测量值的最佳估计,全部不确定成分,包括那些由系统效应,如与修正值、参考计量标准有关联的成分,均贡献于此分散性。

上述定义取自《测量不确定度表示法指南》。】

注意定义附注结尾处明确此术语出自“'93国际指南”。这样带有术语出处的附注在“VIM '93”中是唯一的。它可能出自“VIM '93”对“'93国际指南”权威性的尊重,也可能由于“VIM '93”不想对此词条负责。

“'93国际指南”这样做肯定会给“测量误差和数据处理”领域带来严重混乱,并不利于上述两个国际文件的执行。相当长时期以来,误差评估及数据处理的情况相当混乱。制定一套概念正确清楚,切合实际及简单易行的处理方法的要求甚为迫切。我很久前就意识到,相当一部分混乱来源自国际计量局“INC-1(1980)建议”概念上的混乱和结构上的不足。现在可以肯定“INC-1(1980)建议”体系上存在两个主要缺陷。其一是它没有充分应用统计学和计量学中已经普及的一些概念,加以重新组织和定义,使它们系统化和理论化,而对传统的术语和概念采取否定排斥的态度,引进了一些极不成熟的新概念。例如,“不确定度”、“A类不确定度”、“B类不确定度”、“总不确定度”等等都是这样的概念。必然引起测量误差数据处理领域概念上的长期不稳定和严重混乱。另一个缺陷是它的非数据处理评定方法是一片空白,总是企图将数据处理方法的一些做法强加过来。例如,要求消除误差期望值,要求提供覆盖因子值(往往由于概率分布不确定而是不确定的)及自由度数据等等。一个覆盖面很大的技术文件应该在不同领域允许不同的处理方法,并在不同层次有着不同的要求。强求一致的做法,使得其规定牵强附会、不切合实际、无效及难于执行。

“'93国际指南”作为执行“INC-1(1980)建议”的权威性文件,虽然对一系列问题作了深入研究及改进,但并未能根本消除“INC-1(1980)建议”上述两个体系上的缺陷,相反由于否定“真值”及“误差”,使得其概念混乱变得更为严重。

我国和“VIM '93”及“'93国际指南”对应的规范是JJG 1001-91“通用计量名词及定义”计量技术规范(下文简称“名词规范”)和JJG 1027-91“测量误差及数据处理”计量技术规范(下文简称“误差规范”),长期一直存在要求按国际文件修改我国规范的呼声。目前两个“规范”修订工作正在进行。

1996年中国计量科学研究院(下文简称“计量院”)制定了在该院“试行”的“测量不确定度规范”,我对此“规范”,有两点原则性的保留:

(1) 该“规范”只字未提JJG 1027-91“规范”。工业部门和地方计量机构反映,最不遵守国家技术监督局制定的计量法规的是局本身和计量院。希望局、院能成为遵守局制定的计量法规的模范。

(2) 这“规范”事实上是“'93国际指南”的一个缩写本。它未解决“'93国际指南”遗留下来的任何问题。“'93国际指南”中的问题和错误在这“规范”内同样存在,在一系列问题上和JJG 1027-91“规范”相比明显是倒退的。它显然是无法执行的。

为了给计量院编写一个概念清楚、切合实际和简单易行的贯彻JJG 1027-91“规范”和“'93国际指南”的“实施方法”,1996年我花了接近一年的时间从术语、理论到实例研究了

“'93国际指南”的每一个细节。将研究结论汇总于我编写的“(JJG 1027-91)‘测量误差及数据处理’技术规范及七个国际组织1993年‘测量不确定度表示指南’实施方法(草案)”及其附录中。《中国计量》杂志1997年第3期刊登了“我国的JJG 1027-91‘测量误差及数据处理’技术规范及其解说与1993年七个国际组织的‘测量不确定度表示指南’”一文,初步指出了“'93国际指南”的问题。

1997年5月20日国家技术监督局王以铭副局长主持讨论上述“规范”及“'93国际指南”的座谈会。在会议上我指出了“'93国际指南”四处原则性的概念混乱,以及由于概念混乱造成它的执行必然是混乱的。但其他与会者还是一致强调了和“'93国际指南”进行国际接轨的重要性,相当一部分原因是“'93国际指南”对实际工作的压力。但我同时注意到与会者对“'93国际指南”的理解实际上是混乱的、不符合“'93国际指南”原意的。且无人能否定我提出的问题。

1997年5月25日我参加了在合肥召开的现代误差理论讨论会。在会上我在指出“'93国际指南”上述的四处原则性的概念混乱的同时,还指出“'93国际指南”存在着一个不容回避的遗留问题,两种偏见,三个执行难点。与会者认为问题提得很有道理,但建议从改善“'93国际指南”角度去提这些问题。

由此看来,“'93国际指南”问题最后总得在七个国际组织桌面上才能解决。为此需要进行下列工作:进行必要的学术交流;组织适于向有关国际组织反映的文章并翻译成英文;有组织地向国际组织反映;制定正确的实施方法。我已基本做完了为此事我所能做的工作,与其他同志一起将“JJG 1027-91‘测量误差及数据处理’技术规范及七个国际组织1993年‘测量不确定度表示指南’实施办法(草案)”一文改写成10篇适宜于反映意见的独立文章,其目录如下:

1. 从术语到实例看BIPM, IEC, IFCC, ISO, IUPAC, IUPAP和OIML七个国际组织制定的“测量不确定度表示指南 ISO 1993(E)”中的概念混乱及改善其体系的建议;
2. 误差的随机特性与它区别于其他随机变量的特殊性以及对“测量不确定度表示指南 ISO 1993(E)”的“测量不确定度评定”方法的评论;
3. 工程测量中的实用误差评估方法;
4. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP和OIML七个国际组织制定的“测量不确定度表示指南 ISO 1993(E)”中有关实例的重新评估;
5. 不同测量目的测量列误差评估的完整分析;
6. 工程测量中与“准确度控制”措施密切结合的分层次的误差分解方法;
7. 误差极限值及其直接综合方法;
8. 多次组合测量数据处理的一般方法及其合理性的证明;
9. 与“测量误差”有关基本术语的修改意见;
10. 系统误差的非数据处理(B类)评定方法。

上述10篇文章形成了指明和解决“'93国际指南”问题,能提供一个概念明确正确、切合实际、消除执行障碍、规范化,并与“'93国际指南”兼容的误差评估方法的理论体系。其中,文章1“从术语到实例看BIPM, IEC, IFCC, ISO, IUPAC, IUPAP和OIML七个国际组织制定的‘测量不确定度表示指南 ISO 1993(E)’中的概念混乱及改善其体系的建议混乱及改善其体系的建议”指出“'93国际指南”在“术语定义”、“评定对象”、“分项方法和综合方法”

及“测量数据”四个方面存在原则性的概念混乱。文章2“误差的随机特性与它区别于其他随机变量的特殊性及对‘测量不确定度表示指南 ISO 1993(E)’的‘测量不确定度评定’方法的评论”一文从另一个不同角度指出了“’93国际指南”的“测量不确定度评定”方法还存在着一个不可回避的遗留问题（系统误差的估计及带有系统误差的误差评估），两种偏见（对误差极限值及对系统误差与随机误差的分类），三个执行难点（规范化的完备独立的分项方法，合理的覆盖因子值及自由度）。“’93国际指南”的上述问题都是能排除的。文章3“工程测量中的实用误差评估方法”一文给出了与“’93国际指南”兼容，但排除了其问题的误差评估方法。这种方法概念清楚、处理过程规范、简易而可靠，基本排除了执行过程中的难点，并分有着不同要求的三个评定层次及与实践中的“准确度控制”措施密切配合。

为了保持和“’93国际指南”的兼容性，作了以下工作：推敲“’93国际指南”每一做法的实际意图；对有关术语重新定义；明确不同场合所用术语的对应关系；明确处理过程的关键和解决其难点；回答“’93国际指南”所要求回答的所有问题。可以说上述文章所提供的执行“’93国际指南”的手段不论从概念的明确性方面，或是从切合实际的有效性方面都远远超过了“’93国际指南”本身提供的手段。

文章1和2指出的问题得到了“’93国际指南”本身实例的证实。文章4“BIPM, IEC, IFCC, ISO, IUPAC, IUPAP和OIML七个国际组织制定的‘测量不确定度表示指南 ISO 1993(E)’中有关实例的重新评估”分析了“’93国际指南”中主要实例，包括其正文中三个实例（见其第二节）和其附录H中六个实例（见其第三节）。首先给出实例的出处并引用其给定条件及处理结果。然后再按该文规定的（与文章3的规定相同）方法重新评估。然后比较两个处理过程进行评述。九例中“’93国际指南”附录H中的H.3和H.5两例是纯数据处理的实例。在这两例中的两种处理结果完全一致。该文没有引入这两例的分析。其他七例涉及“非数据处理”评定的领域，“’93国际指南”的规定就显示出其明显的缺陷。其评定结果都存在不同的问题而且是不可信的。

文章9“与‘测量误差’有关基本术语的修改意见”一文是根据对与“误差”有关术语的全面分析提出的对“VIM ’93”的修改建议。分析表明，“VIM ’93”术语内容基本准确，体系基本完备，是国际上统一计量学术语的良好基础。但鉴于上述情况的存在，“VIM ’93”还是有必要作一定的局部修改和补充，其原因如下：

(1) 对“’93国际指南”的非议不宜置之不理，它引入的混乱必须消除。

对“不确定度指南”的非议不宜置之不理，首先在它的广泛传播及权威性。受到它所非议的术语在应用上必然出现障碍。例如术语“[测量]误差”的使用就遇到强烈地抵制。受到非议的术语又是最基本的，在计量信息交流中是无法回避的。其次“’93国际指南”的非议是有其理由的，虽然这些理由出自“认识误区”不能成立，但这些“认识误区”却是经常困惑人们的。因此，受到非议的有关术语有必要增加排除非议的说明。

在“’93国际指南”的影响下，“VIM ’93”的某些术语必然存在某种的概念混乱。其(3.09)“测量不确定度”就是一条言不达意，概念混乱的术语，它必须重新定义。受较小影响的还有(1.20)[量的]约定真值，(3.5)测量准确度等，它们应表述得更符合实际情况及明确。

(2) 误差评估是计量学定量分析的理论基础，在计量学中占有重要地位。“VIM ’93”几乎没有列入和误差评估有关的术语，需要作必要的补充。但其中部分术语（如标准差、覆盖因子、自由度等）表述的都是随机变量共有的特性，因此应该引用统计学通用术语。这样就

有必要引入一个和“计量”有关的“统计学术语”的附录，将有关术语安置在这一附录内。文章9给出了这样的附录，其主要术语取自我国GB/T 3358.1—93国家标准《统计学术语 第一部分 一般统计学术语》（下文简称“统计学术语”）；为适应“计量”工作（主要是误差评估）的需要，在内容上作了必要的调整。引入“有关的统计学术语”的附录对“VIM '93”其他的术语的表达也有好处，采用统计学通用术语可以将与随机特性有关的术语（如术语系统误差与随机误差）表达得扼要而明确。

有关量值的定义的最准确的表述方法是用数学公式表述，如“测量误差是测量结果减去被测量真值之差”等。严格说，不能用数学公式表示的有关量值的定义往往留有含糊不明确之处。一个全部用数学公式表述的术语体系就有一个协调一致的术语表示符号体系的要求。这样的符号体系的建立要求较高的水平，因此有一个建议性符号体系有利于交流信息，但不宜强求统一，因此以建议性的列入附录为好。这里，“'93国际指南”的经验是值得借鉴的，它在附录C中引用了统计学术语，在附录J中表述了其符号体系。显然在它的分析和信息交流过程中感到了需要这两个附录。

事实上对“'93国际指南”的全面的研究工作是从研究有关术语开始的，不弄清楚“'93国际指南”的有关术语是无法明白其内容的实质和涵义的，没有合适的术语体系就无法表达正确的技术思想，而“'93国际指南”的问题正是从术语概念混乱开始的。

文章5~8是文章1~4的理论基础，其内容是误差理论中最重要却长期悬而未决的问题。

“'93国际指南”的一系列规定的依据是数据处理方法，而最基本和常用的数据处理方法是等精度多次测量列的数据处理。文章5“不同测量目的测量列误差评估的完整分析”指出，同一形式的测量列可以用于完全不同的测量目的，采用不同定义的“被测量真值”，因此有着不同的“误差”和不确定度（即误差统计特征估计值）。根据这一分析文章1指出如果“被测量真值”和“误差”定义不明，不确定度的评定的研究对象就是不明的，这样的评定是在算“糊涂帐”。“'93国际指南”对文章4的第2—3节的例（“'93国际指南”正文4.3.7款中的例2及5.1.5款中的例：数字电压表测量电压）的不确定度的评定就是这样的一笔“糊涂帐”。文章4在其第2—3节中对此例与测量有关的三种误差的极限估计值（即扩展不确定度）分别作了评定。文章1根据上述分析指出，“'93国际指南”附录D否定“被测量真值”和“误差”的理由是不能成立的，意见是不可取的。这一附录的内容实际上是“'93国际指南”一些重要概念混乱的根源，同时也是国内外误差评估领域概念混乱的重要根源。

文章5同时指出，测量列的数据是否有效，取决其测量条件是否符合测量目的。测量条件不明的测量列的处理结果的意义是不明的。并指明了为实现给定的测量目的控制测量条件的主要原则。在“'93国际指南”中从未提出要提供测量数据测量条件的要求，其实例中，也从不给出所处理数据的测量条件。因此文章1认为“'93国际指南”的测量数据是混乱和意义不明的。

文章5还指出，测量列的数据处理结果反映了其测量条件下所有因素所引起的测量结果的全部波动，但无法提供这个测量条件下误差的条件期望值的任何信息。其分析进一步表明，这个条件期望值的内容是测量时所用标准测量设备的系统误差值，这个误差的唯一的确定方法是对这标准测量设备进行检定（或校准）。当上述标准测量设备的示值用其检定确定的修正值修正后，残留的（即未定的）系统误差值将由确定修正值的检定误差及由检定到测量这一时间区间内标准测量设备系统误差的变化值组成。测量是通过确定这一误差进行“量值溯

源”的。这误差是量值溯源（或量值传递）过程唯一得到传播的误差，也是对测量误差采用正确测量条件的测量列数据处理（即A类）方法评定后唯一需要补充确定的误差。它的确定是“INC-1（1980）建议”和“’93国际指南”必须求助非数据处理（即B类）评定方法的根本原因。未定系统误差是误差期望值的一部分，有着确定的值，其评估必须求助量值溯源（或量值传递）过程，它的确定值无法知道，仅能用其可能值估计它。其可能值事实上是扩大了观察范围的随机变量。因此未定系统误差的实际评估是按随机变量处理的，这时评估的已是它的可能值，而非其确切的值。

在用A类方法对误差的中心化变量（即随机误差）作出正确评估后，“INC-1（1980）建议”要求用B类方法作补充评定，正确评定对象应该局限于误差期望值（即系统误差），否则随机误差就被重复计算了。一般情况下，这样补充评定的要求是合理的。但如果能确定系统误差是可以忽略的，则补充B类方法评定就是多余的。文章4的3—1例（“’93国际指南”附录H的H1例：端度规块校准）中比长仪误差就是这样的例子。

特别要指出，大多数误差项的B类评定结果是包括其随机误差的，在这种情况下并不需要补充A类评定。但如果这时A类评定结果小于B类评定结果，则将是B类评定结果可靠的重要旁证。综合时应仅采用B类评定结果，不应再将A类评定结果引入综合。

“系统误差”和“随机误差”是一种完备独立的分类方法，两类误差的性质，评估方法有明显的差别，其交界交叉区“未定系统误差”的组成也是明确的，通过“准确度控制”措施中“标准测量设备”的检定（或校准）它是能被评定的。因此这样的分类方法完全有其存在的理由。相反A类不确定度与B类不确定度的分类方法既不完备，又不独立，无法保证主要误差项不被遗漏，又是误差项多次重复的理论根源。“’93国际指南”坚持用A类不确定度与B类不确定度的分类方法取代“系统误差”和“随机误差”分类方法是文章2指出的它的两种偏见之一。

在分项方法概念上混乱和缺乏有效的、规范化的、完备独立的分项方法是执行“’93国际指南”最主要的障碍。分项实际上是对误差按其产生原因进行的，不确定度分项本身就是一种混乱的概念。误差分解的最重要的原则是“主要误差项不得遗漏和重复估计”。“’93国际指南”在其4.1条“测量模型化”中建议用“数学模型”来指导分项，这工作要求对测量原理有全面和高度地掌握。在“’93国际指南”的四个有分项的实例中，文章4的3—2例和3—3例（“’93国际指南”附录H的H.2例和H.4两例）中“’93国际指南”的分析是遗漏了或基本没有评估其主要误差项，因此这样的评定根本是无效的；而在3—1例（“’93国际指南”附录H的H.1例）中“’93国际指南”对比长仪基本误差作了多次重复估计，使得它的比长仪基本误差标准差的评估值超过了正确评估的比长仪基本误差极限值。相反，比长仪示值温度附加误差在表H.1中的对应项明显偏小，在表H.1中序号4的误差项显然就是比长仪示值温度附加误差，采用0为传播系数是错误的，正确的传播系数应该是 $\sigma(\delta\alpha)$ 。表H.1中没有反映出 $\delta\alpha$ 的中心化变量与 0 的中心化变量共同作用所引起的误差项，它是比长仪示值温度附加误差主要组成项。在3—4例（“’93国际指南”附录H的H.6例）中同样存在误差项多次重复估计的问题。在“’93国际指南”本身的实例中，分项尚且这样漏洞百出，可以想象在推广执行时会出多少问题。在这些实例中引起误差项的重复评估错误的另一个理论根源是“INC-1（1980）建议”的错误概念。它认为一个误差项不同方法估计的特征值是相互独立的“不确定度”，并建议同时进入综合。因此文章1指出“’93国际指南”在分项及综合方法上存在严重的概念混乱。

文章6“工程测量中与‘准确度控制’措施密切结合的分层次的误差分解方法”在对测量实际过程仔细严密的基础上,结合广泛存在的“准确度控制”措施,综合各种现行的误差分解方法,给出了规范化的、有三个不同要求层次的误差分解方法。由于这个分解方法是和“准确度控制”措施密切结合的,其较低的两个层次并不要求执行者对测量原理有很深的了解。基本解决了误差评估中最主要的难点。

将所有误差评估值归算为标准差是“INC-1(1980)建议”和“’93国际指南”不确定度评定的重要基点。当随机变量有着确定的概率分布时,变量的标准差也有着明确的数学定义和严格的运算法则。因此,它成为数理统计中表示分散性的主要统计特征值。在用数据处理方法评估误差时,确定的测量数据当然有着确定的概率分布,因此在数据处理中采用标准差作为误差的主要统计特征值是合适的。但从“准确度控制”的需要和方便考虑,还是希望提供高概率误差极限值,这是“INC-1(1980)建议”和“’93国际指南”规定“扩展不确定度”的原因,并由此出现了合理“覆盖因子”的问题。当误差用非数据处理方法评估时,情况就不同了。这时有关系的变量都受到“准确度控制”的制约。这控制通常是用删除被控制变量超过其允许值的异常值的方式实现的,这样的控制可称为极值控制。受极值控制的变量有着明确可靠的极限值,而其概率分布依赖于控制变量的实现条件而是不确定的,并由此其“标准差”与“覆盖因子”也都是不确定的。在这条件下将“极限值”归算成“标准差”是不合适的。坚持这种“归算”要求是出自对“极限值”的偏见。这种偏见可归纳成以下两点:

(1) 认为“标准差”是可相比的,而“极限值”由于“覆盖因子”的不同,是不可比的。

(2) 认为统计学为标准差的综合提供了理论严密的综合方法,而对于“极限值”则缺乏类似的综合方法。

如果认为误差统计特征值的大小是表征着测量结果逼近被测量真值的可靠程度,则按一定的可靠性水平的可靠性指标确定的“极限值”是可比的,“极限值”越大,可靠程度越低。相反,“标准差”是不可比的,因为不同概率分布的“标准差”代表着不同的可靠程度。

统计学在组成项方差值确定及方差估计值为其期望值的条件下,给出了严格方差综合公式。受“准确度控制”的误差项的方差是不确定的,其估计值通常是其上限值。在这种条件下严格的方差运算法则也无法给出明确可靠的误差标准差的估计值,因为被估计量值本身是不确定的。相反,在方差严格运算法则基础上,是可以推导出普遍适用的,统计学上合理的“中心化极限值”(即“扩展不确定度”)直接综合法则。文章7“误差极限值及其直接综合方法”一文对极限值的定义、可靠性指标与可靠性水平作了全面的研究,采用“峰度”作为数学工具将可能出现的误差分布为三类,确定了研究采用的三种典型分布,并大致估计了其出现概率的数量级。在这基础上证实了方和根法对独立中心化极限值综合的普遍适用性。为极限值的直接综合奠定了理论基础。“极限值”的直接综合回避了用非数据处理方法确定“极限值”的各组成项概率分布及“覆盖因子”的研究。“覆盖因子”合理值的确定是执行“’93国际指南”的第二个难点,根本原因在于非数据处理方法确定“极限值”的各组成项没有确定的概率分布及“覆盖因子”。“’93国际指南”武断地规定分布不明的误差项的极限值归算成标准差时“覆盖因子”值采用为 $\sqrt{3}$ 。由于受“准确度控制”的误差项极限值的可靠性取决于控制措施删除异常值的准确性和有效性,其可靠性指标(包括“覆盖因子”)通常是不清楚的,其概率分布又不确定的,因此对这样的误差项归算时采用 $\sqrt{3}$ 作为“覆盖因子”值的规定几乎对全部误差项都有效。“’93国际指南”又建议根据标准差确定“扩展不确定度”(即

中心化极限值)时,“覆盖因子”值采用2~3。这样一个来回折算就无缘无故地将所有的非数据处理方法确定“极限值”扩大了 $(2\sim 3)/\sqrt{3}$ 倍,这样的来回折算显然是不合理的。因此在存在误差项B类评定结果时,采用“极限值”的直接综合方法为宜。这样的综合简单易行,而结果可靠且一致性高。

文章7对合理的“覆盖因子”值进行了全面的研究,提出采用“覆盖因子”的“峰度”线性表达式替代“显著性水平”定义目前统计学中通用的“一般异常”和“高度异常”两个约定可靠性水平的建议。并建议在情况不明时取2作为“覆盖因子”的约定值。

文章7指出“自由度”在“数据处理”中,有着明确的数学定义,很容易确定,可以用来判断处理结果的可靠程度。将“自由度”推广到“非数据处理”领域就失去了原有的自然和明确的数学概念,变成牵强附会,无用及导向错误判断的概念。并建议在情况不明时取10作为“自由度”的约定值。

注意文章7在讨论过程中,误差是允许带有期望值的,并其分布是不限的。

文章8“多次组合测量数据处理的一般方法及其合理性的证明”以适用面最广,最结合“INC-1(1980)建议”和“’93国际指南”规定的形式,在简单明确的前题下,讨论了通用的最小二乘方法的数据处理过程及其合理性。讨论过程对误差的概率分布未作任何限制,因此其处理方法对任何分布的误差都是普遍适用的,并不局限于对正态分布有效。常用的数据处理方法(如等精度测量列的平均值法、不等精度测量列的计权平均值法、相关测量列的线性回归等)都是最小二乘方法的特例。它们的处理结果同样对任何误差分布都有效。

文章10全面地讨论了系统误差的确定和消除方法,弥补了“’93国际指南”所遗缺的,而又是“误差评估”必不可少的内容。

上述10篇文章是为向BIPM等七个国际组织反映“’93国际指南”的问题而编写的。现在首先在国内发表,是基于下列目的:

(1)编者希望10篇文章的内容首先在国内得到尽可能多的认同。这些文章提出的问题、解决办法及理论对于大多数有关人员来说是新的或不十分清楚的,认同这些观点需要一个交流传播的过程。希望这些文章的发表能加速这一交流传播的过程。

(2)误差理论的关键问题是理论联系实际。“’93国际指南”的问题就出在概念混乱及其研究基点与实际不符。同样10篇文章的内容也应该经得起实践的检验和推敲。目前国内外正在积极推行“’93国际指南”,而由于其本身的原因,肯定会出现混乱及难点,这时请研究上述各文指出的问题及解决办法,应该对消除混乱和解决难点有所帮助。在工程测量中可以试用“工程测量中的实用误差评估方法”一文中提出的方法对误差进行评估。在进行这些实践的同时,也希望对10篇文章中不妥之处提出异议及改进意见,使其内容能更趋完善。

10篇文章是以“’93国际指南”有文字表达的内容和实例及测量实践的实际情况作为依据指出其具体不妥之处、不妥的原因及解决办法的。希望对10篇文章内容的异议也能以其有文字表达的内容及测量实践的实际情况作为依据,指出其具体不妥之处及解决办法,不要用简单的否定或肯定,或以某种需要作为藉口来搪塞敷衍。

常见的异议之一是“误差难以弄清,不确定度是一种实用的人为约定,不宜深究”。

文章2明确指出:“测量误差”作为“随机变量”没有确定的大小,可以也仅能用其统计特征值来表示其统计意义上的大小。这样的特征值主要有:误差期望值、误差标准差、误差极限值、误差有效值(均方根值)及误差中心化极限值。误差评估就是确定这些特征值的估

计值。而“测量不确定度”事实就是表示误差随机部分（即分散性）大小的统计特征估计值。因此“测量不确定度”评定仅是误差评估的一部分。该文又指出，误差与被测量值相比是一阶微量，它表征着“测量结果”作为“被测量真值”估计值的可靠性，对它的评估的确“不宜深究”。该文明确了误差的二次以上项（即误差的误差）在一般误差评估中是可以忽略的，仅在评定评估方法的可靠性时才需要研究它们。又指出，除用于误差修正值的误差期望估计值外，其他误差统计特征值采用其绝对值足够可靠的上限值作为其估计值，而不企求准确地确定这些特征值的期望值。明确了“不宜深究”的界线。在传统的误差理论中，还可以用“微小误差准则”来明确“不宜深究”的范围。

但“不确定度”术语的定义，被评定“不确定度”的误差涵义，及“主要误差项不得被遗漏及重复估计”的原则显然都是不属于“不宜深究”的范围的。不清楚什么是“不确定度”，不清楚评定的是什么“误差”的“不确定度”，这样的“不确定度评定”必然是混乱的。“主要误差项被遗漏及重复估计”的误差评估结果是不可信的。“’93国际指南”的混乱正是出现在这些必须深究之处。相反，“’93国际指南”要求提供的“覆盖因子”和“自由度”是属于“不宜深究”的研究的“误差的误差”的范畴，“’93国际指南”的规定是“拣了芝麻，丢了西瓜”。

一个误差的两种评估结果差异到两者之比到1.2是正常的，到1.4还可以不加深究。但对于“’93国际指南”的实例，或其某些误差项“不确定度”的评定结果比正确评估值可能小一个数量级（如其附录H的H.2和H.4例中未估计最主要误差项），也可能是大到一倍以上（由于采用不合理的“覆盖因子”值进行换算及其组成项的重复计算），在这样的差异情况下再提“不宜深究”是不适宜的。

误差所以无法弄清，原因是对“测量准确度的实际控制方法”缺乏应有和必要的研究。

“INC-1(1980)建议”的制定者所以无法处理“未定系统误差”，原因是对测量所用标准测量设备的检定过程缺乏应有的研究，而拘泥于一般的数据处理方法。“’93国际指南”所以存在如此多问题，原因之一是它的B类方法为一片空白，对广泛存在并控制着误差的“准确度控制”措施体系熟视无睹，视而不见，而总是企图将A类方法的做法强加给B类方法。

另一个不同意见是强调和国际上保持一致的重要性。这不是一种学术上的意见，而是一种出自实际利益考虑的意见。实际利益无疑是重要的，在国际规定的问题是次要和可以回避时，不妨为了实际利益将次要分歧放在一边。但实际利益终究无法将错误变成正确。“’93国际指南”的问题是如此根本，它的认真执行必然引起混乱。在这样的情况下，明确问题，作出正确的规定，并积极向有关国际组织反映，以求国际上能在正确的基础上达到新的一致，是唯一的解决办法。当然为了减少过渡期间的混乱，上文已指出了我们与“’93国际指南”兼容的努力。这正是我们在进行的工作。这工作在国内外国会遇到很大的阻力，没有各方面的支持，特别是主管计量及质量管理行政部门的支持，工作是很难见效的。在这里我们期望我们的努力能得到必要的支持。

一、从术语到实例看 BIPM,IEC,IFCC,ISO,IUPAC,IUPAP 和OIML七个国际组织制定的 “测量不确定度表示指南 ISO 1993(E)”中的 概念混乱及改善其体系的建议

钱钟泰 宋明顺 顾龙方 邹本霞

提要:

本文指出“测量不确定度表示指南ISO 1993(E)”(下文简称“’93国际指南”)中至少有四处的基本概念存在原则性的问题:

(1)“测量不确定度”是一个言不达意,概念混乱的术语。术语定义字面上的涵义肯定不是术语主要起草人所表达的意思。

(2)由于“’93国际指南”对“不确定度”缺乏明确的“误差”和“被测量真值”的概念。因此,它的评定对象是混乱的,它的“不确定度”评估经常是在算“糊涂帐”。

(3)“’93国际指南”把“误差分项”混淆成“不确定度分项”,而以“不确定度”的评估方法作为分项的依据根本无法保证主要误差项不被遗漏。“’93国际指南”建议一个误差项的不同评估方法确定的“不确定度”同时进入综合是一种误差项重复评估的错误。由于这一理论上的谬误在“’93国际指南”评估“不确定度”的实例中是主要误差项的遗漏和重复评估并存,大部分评估结果是不可信的。

(4)测量列的数据处理结果的意义及可靠程度取决于测量列测量条件的控制是否符合研究目的。“’93国际指南”未提出交待测量列测量条件的要求,在其实例中也从未交待过测量列测量条件。因此它的评估结果的实际意义及可靠程度都是存疑的。

上述结论是将“’93国际指南”的术语定义及基本概念和“不同测量目的测量列误差评估的完整分析”一文的结果相比较得出的,并得到对“’93国际指南”主要实例重新分析的结果的支持。在“’93国际指南”的九个主要实例仅有两个纯数据处理实例是正确的。其他七个和测量有关的实例都由于上述概念混乱存在不同程度的严重问题,其评估结果是不可信的。

在本文的最后提出了改善“’93国际指南”体系的建议。

目前国内外误差评估及数据处理领域的情况相当混乱。我们很久前就意识到,相当一部分混乱来源自国际计量局“INC-1(1980)建议”(见文献3的07款,下文简称“’80国际建议”)概念上的混乱和结构上的不足。1993年七个国际组织[BIPM(国际计量局),IEC(国际电工委员会),IFCC(国际临床化学联合会),ISO(国际标准化组织),IUPAC(国际理论与应用化学联合会),IUPAP(国际理论与应用物理联合会)和OIML(国际法制计量组织)]制定了“测量不确定度表示指南ISO 1993(E)”(文献3,下文简称“’93国际指南”)作为执行“INC-1(1980)建议”的权威性文件。为了研究如何执行“’93国际指南”,我们从术语、理

论到实例研究了“’93国际指南”的每一个细节。现在可以明确无误地指出“INC-1(1980)建议”和“’93国际指南”主要问题的所在。本文将根据对“’93国际指南”的术语和实例的分析所暴露出来问题,说明存在于“’93国际指南”中的概念混乱,并最后提出了改善其体系的建议。

(一) “’93国际指南”的基本术语和其概念上的混乱

1-1 基本情况

“’93国际指南”的最主要的术语显然是“测量不确定度”。它对这术语的定义如下:

【3.9 测量不确定度

与测量结果相关的参数,表征合理地赋予被测量的值的分散性

注:

1. 此参数可以是,如标准差(或其倍数),或置信区间的半宽度。
2. 测量不确定度一般由许多成分组成。一些成分可以由测量列结果统计分布估计,由实验标准差表征。另一些也可用标准差表征的成分是基于实验或其他信息的概率分布来估计。
3. 应这样理解测量结果。它是被测量值的最佳估计,全部不确定成分,包括那些由系统效应,如与修正值,参考计量标准有关联的成分,均贡献于此分散性。

上述定义取自《测量不确定度表示法指南》^[10]。】

上述定义取自相同七个国际组织在1993年制定的另一个文件“国际通用计量学基本术语”(第二版)(文献5,下文简称“VIM ’93”)。注意定义附注结尾处明确此术语出自“’93国际指南”。这样带有术语出处的附注在“VIM ’93”中是唯一的。它可能出自“VIM ’93”对“’93国际指南”权威性的尊重,也可能由于“VIM ’93”不想对此词条负责。

“’93国际指南”对“VIM ’93”就显得不那么尊重。“’93国际指南”在它的附录B中引用“VIM ’93”的有关术语时,加了多达九处的“指南评注”。评注的核心是否定“VIM ’93”的“[量的]真值”及“[测量]误差”的词条。“’93国际指南”还在它的以“真值、误差和不确定度”标题的附录D中表述了下列观点。

在其附录D一开始就指出,指南的【不确定度概念是基于测量结果及其评定不确定度,而不基于不能知道的“真”值及误差】。

在其D.3.5款中的内容如下:

【D.3.5 术语“被测量真值”或量的真值(常简称真值)在本指南中避免使用,因为“真”字视为多余……】

在这里,它认为被测量“真”值就是被测量值。而却在D.2款中引入了“认得量”的词条:

【D.2 认得量

D.2.1理想地,测量得的量应与被测量定义相一致,后者经常不能认识,而测量是对被测量的近似量进行。】

而在附录D的D.4款中误差定义如下:

【D.4 误差

由于观测的随机变化(随机效应),确定系数效应修正的不完善,一些物理现象(也是系统效应)知识的不完善,故认得量的测量是不完善的。已修正结果并不是被测量值,即存在误差。……】

将这些定义和“VIM '93”的有关内容相比较,这里所谓的“认得量”就是“被测量真值”。在这样“'93国际指南”实际上是废掉了“被测量真值”的老的术语,而起用了“认得量”的新术语,而这一切都是在含糊不清的定义下进行的。

“'93国际指南”这样做肯定会给“测量误差和数据处理”领域带来严重混乱,并不利于上述两个国际文件的执行。就是在这种背景下我们分析了上述两个权威性国际文件的意见分歧。分析结果表明,“'93国际指南”中至少有四处基本概念存在原则性的问题:

(1) “测量不确定度”是一条言不达意,概念混乱的术语。这一术语的主要起草人都不清楚自己在说些什么。术语定义字面上的涵义肯定不是术语主要起草人所想表达的意思。

(2) 当采用不同的“被测量真值”定义时相同的测量结果可以有着不同定义的“误差”。不同的“误差”当然有不同的统计特征估计值(即“不确定度”)。由于“'93国际指南”对“不确定度”缺乏明确的“误差”和“被测量真值”的概念,因此它的评定对象是混乱的,其“不确定度”评估经常是在算“糊涂帐”。

(3) “'93国际指南”把“误差分项”混淆“不确定度分项”,而以“不确定度”的评估方法作为分项的依据根本无法保证“主要误差项”不被遗漏。“'93国际指南”建议一个误差项的不同评估方法确定的“不确定度”同时进入综合是一种误差重复评估的错误。由于这一理论上的谬误在“'93国际指南”评估“不确定度”的实例中是“主要误差项的遗漏和重复评估”并存,大部分评估结果是不可信的。

(4) 测量列的数据处理结果的意义及可靠程度取决于测量列测量条件的控制是否符合研究目的。“'93国际指南”未提出交待测量列测量条件的要求,在其实例中也从未交待过测量列测量条件。因此它的评估结果的实际意义及可靠程度都是存疑的。

1-2 不同测量目的测量列误差评估的完整分析

为说明上述判断的根据,本书文章五对同一测量列的不同测量目的的误差评估作了全面、详细和严谨的分析。

如果测量列 $Y_i(i=1\sim m)$ 的测量结果示值用 $Y(\tau)$ 表示,其被测量值用 $Y_0(\tau)$ 表示及测量系统示值误差用 $\Delta Y(\tau)$ 表示,则有:

$$Y(\tau) = Y_0(\tau) + \Delta Y(\tau) \quad (1-1)$$

当被测量值 $Y_0(\tau)$ 及测量误差 $\Delta Y(\tau)$ 是随机变量时,测量结果 $Y(\tau)$ 也将是随机变量。在时间 $t = \tau$ 时对随机变量 $Y(\tau)$ 作 m 次独立观察,得出测量列:

$$Y_i \quad (i=1\sim m) \quad (1-2)$$

用 $E(X)$ 表示随机变量 X 的期望值。用 X_c 表示随机变量 X 的中心化变量,其定义为:

$$X_c = X - E(X) \quad (1-3)$$

中心化变量 X_c 反映了变量 X 的全部随机性,变量 X 及其中心化变量 X_c 作为随机变量没有确定的量值,其大小仅能用其统计特征值表示。在这里称表示随机变量 X 某种随机特性的参数为变量 X 的统计参数,如期望值,方差,矩及峰度等;称表示变量 X 大小的与变量 X 同量纲的统计参数为变量 X 的统计特征值,如期望值 $E(X)$, 均方根值(有效值) $\sigma_0(X)$, 极限值 $U_0(X)$, 标准差 $\sigma(X)$ 及中心化极限值 $U(X)$ 等。其中,均方根值(有效值) $\sigma_0(X)$ 和标准差 $\sigma(X)$ 的定义如下:

$$\sigma_0(X) = E(X^2)^{1/2} \quad (1-4)$$