

测量误差 与不确定度评估

林洪桦 编著

CELIANG WUCHA
YU BUQUEDINGDU PINGGU

 机械工业出版社
CHINA MACHINE PRESS



测量误差与不确定度评估

林洪桦 编著



机械工业出版社

本书以应用“测量不确定度指南”ISO 1993 (E) 及其扩展应用为纵线, 全面、系统地论述了测量结果及其不确定度评估与误差理论及其分析方法的密切关系、基本理论与基本方法。尤其是集成地阐述了尚未包含于该“指南”中的误差分析的各种现代方法(如各种最小距离准则、稳健性、熵分析和优化、Bayes 估计、数值仿真分析、支持向量机等方法), 并简要地述及变量动态测量的分析方法。本书既具有实用价值, 又有较高学术性。

本书适合各行各业从事测量误差分析及不确定度评估的人员、高等学校教师及学生阅读, 并可作为研究生教材。

图书在版编目 (CIP) 数据

测量误差与不确定度评估/林洪桦编著. —北京: 机械工业出版社, 2009. 8

ISBN 978-7-111-27864-1

I. 测… II. 林… III. ①测量误差②测量-不确定度 IV. P207 TB9

中国版本图书馆 CIP 数据核字 (2009) 第 128463 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 林春泉 责任编辑: 赵玲丽 责任校对: 陈延翔

封面设计: 鞠 杨 责任印制: 杨 曦

保定市 中画美凯印刷有限公司印刷

2010 年 1 月第 1 版第 1 次印刷

184mm × 260mm · 30.5 印张 · 758 千字

0001—3000 册

标准书号: ISBN 978-7-111-27864-1

定价: 78.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

社服务中心: (010) 88361066

门户网: <http://www.cmpbook.com>

销售一部: (010) 68326294

教材网: <http://www.cmpedu.com>

销售二部: (010) 88379649

读者服务部: (010) 68993821

封面无防伪标均为盗版

前 言

科学规律的形成、新技术和新产品的开发，都离不开实验与测试的分析和总结，而实验与测试又总不免存在误差、不稳定和不确定等非本质因素的影响。只有通过误差分析及合理的实验数据处理，尽量分离出非本质因素的影响，表示出合乎实际的最佳结果，且评定其尚存的不确定度，才有利于总结出科学规律及开发新技术、新产品。这已是孰知的公理，也是人们不断研究误差理论及运用误差分析与数据处理方法的动力和目的，在国内外均已有了体系较完整的误差理论及数据处理方面的教科书与学术专著。

然而，人类的认识总不免有局限性，且科学发展又有其阶段性，误差理论与数据处理这一学科自然也不例外，又因其所研究的对象多为不确定性因素，而使之更加复杂。在实际应用中也存在不一致、不规范等问题。

首先，误差的定义是相对于无误差的真值作出的，而真值恰为待求的未知量，又因误差难免而不可确知。这就使得误差不便于按定义直接定量表示，在以往的实际应用中，出现了多种不同的表示方法，如极限误差、中误差、平均误差等等。

其次，误差是按其性质主要分为：系统误差和随机误差，且随着误差性质的不同，相应的处理方法也就各异。如：对于系统误差，应尽量发现它，并加以修正。然而，实际上未必能发现其存在，并已知其修正量。况且即便已知该修正量，还需估计该修正量的误差。对于随机误差，则通过多次重复测量，取平均而减少其影响。这样就存在估算均值误差的问题。进而有时需对这些误差再予细分，从而又出现更多各不统一的术语与各不一致的处理方法。如以往曾有：未知系统误差、不定系统误差、可变系统误差、随机系统误差、半系统误差等含义不确切的术语。

在误差合成方面，以往由于对误差分类和表示不一致，导致出现部分按最大误差合成、按绝对和合成、按平方和合成等互不一致的方法。

此外，对于粗大误差判别及剔除异常数据的方法，多局限于正态性数据及逐个判别剔除方法，而判别界限及剔除方法也很不一致。

因此，历来都存在对误差的表述、分析、处理等不一致、不统一、不规范的问题。

总之，上述问题导致在表达测量结果及其误差上出现不少差异，尽管以往的那些误差表述的主张，多数不无道理，却影响了科技交流及成果评价，以及产品质量评定与互比。为了统一和规范实验与测试中非本质因素影响的表述，国际计量局（BIPM）及国际标准化组织（ISO）等七个国际组织联合公布了“测量不确定度指南”ISO 1993（E）（本书简称“指南”）。“指南”目的在于：促进以足够完整的信息表述测量不确定度；提供测量结果国际可比对的基础；归结出广泛可接受的测量不确定度评定与表示的一般原则及通用规则；既采用了公认已成熟的、传统的误差分析与数据处理方法，也应用了某些现代方法，并强调其可操作性与实用性。

以上所述，正是本书以应用“指南”为纵线来论述误差理论及其分析方法与数据处理的初衷。

尽管“指南”总结出了各国专家公认成熟的评定测量不确定度的见解及其估算与表示方法，并集中反映了当前国际上的新研究成果。却也指出其前提条件和未说细、未说完的问题。如：其所指的测量结果多是已修正者，而未谈及修正方法；不打算在不确定度度量中考考虑数据会有失误值或异常值，即未给出剔除异常数据或排除粗大误差影响等方法；即便对 A 类评定的论述也并未谈全，即还可补充许多更为复杂情况下的统计处理方法；强调仅提供了评估不确定度的框架，不能替代更周密的思考、专业的分析技巧；指出可推论至动态测量不确定度评定，却未具体谈及等。实际上，在“指南”的实施后，也已发现存在待补充与扩展的问题。

尽管该“指南”的贯彻实施已有十几年，却因其难免存在的局限性，如未含对系统误差及其修正量的处理、排除粗大误差的影响、量仪检定中评定准确度的特殊要求以及各种特殊情况下新型数据处理方法的应用等等。至今还存在对“指南”有异议或不同见解者。

尤其，随着科技的不断发展，对测量可靠性与准确度有更高的要求，促使误差分析及数据处理中出现了许多新观点及新方法，并已成为该学科的研究重点课题及发展动向。其中许多均为当前科技发展中的新分支的应用。如：

从传统正态性误差进一步深入研究而扩展至非正态分布误差，尤其是非对称分布误差等研究。包括各种误差分布类型的表示方法、评定指标、估算方法、合成方法等各方面的探讨、研究及应用。

从传统上基于最小二乘法的统计处理扩展至适应于各种专业领域的不同技术要求和不同条件下的其他最小范数或最小距离准则的统计处理方法。包括残差绝对值之和为最小的最小一乘法，残差最大值为最小的最小 α 乘法，或两者之间的最小 m 乘法等。也包括最小二乘法最佳估计的假定条件不满足或有所偏离时，而采用的稳健统计 (robust statistic) 处理方法；充分利用已有的知识、经验、资料等先验信息与现有数据一起采用 Bayes 统计处理方法；按最大熵准则和最小互熵准则的所谓熵优化原则 (entropy optimization principle) 的处理方法；按统计学习理论 (statistical leaning theory) 形成的支持向量机方法；以及模糊推理分析的处理方法；灰色系统分析的处理方法等等。

从常量重复测量数据传统的统计处理方法扩展至非统计的数据处理方法，又从其成批算法 (batch algorithm) 扩展至变量测量过程的动态数据处理方法，尤其是长过程测量中 (包括常量长时间重复测量过程) 数据统计特征可变的递推算法 (recursive algorithm) 与移动算法 (moving algorithm)。某些数据处理的新方法多数是在 20 世纪七八十年代后才发展至逐步成熟和实用化，可统称之为现代方法。

由于上述误差分析及数据处理中存在许多固有的问题、尚未完善解决的难题，及发展中不免会产生一些新问题。在“指南”公布与实施后，在一定范围内起到了规范化表示测量结果及其不确定度的作用。又存在对“指南”的补充与扩展应用问题。

因而本书曾拟副标题为：兼论“指南”的扩展应用。

鉴于上述情况与需要，结合作者长期从事“误差理论与数据处理”的教学经验与科研成果，以及学习与应用 ISO 1993 (E) “指南”的心得和体会，予以归纳和总结而撰写出本书，在本书中作者力图重点阐明：误差理论与测量不确定度评定的紧密关系；测量误差分析与不确定度评估的现代方法；“指南”待补充和扩展应用的内容，有益于正确理解及合理实施及扩展应用该“指南”。

全书内容共分为 12 章：

第 1 章概述了测量不确定度表述的意义和作用，同时简要地分析了以往误差分析中存在的问题及误差理论发展的现况，以及 INC-1 (1980) 建议书的提出及其导致发布 ISO 1993 (E) “指南”的过程与当前推广应用的情况。

第 2 章主要阐述了与测量不确定度密切相关的基本概念和术语及定义，并以“国际通用计量学基本术语”（第 2 版）中的有关术语的定义为主，与“指南”所论述的基本概念作比较，以加深对其理解。同时着重论述了测量不确定度与误差理论的关系及对“指南”的正确理解与扩展应用。

第 3 章着重分析了不同被测量及不同影响量所影响下的测量数据的特点，包括其组成、数学模型以及相应的各种数据处理方法等。并举出其典型示例。同时，简述了数据处理方法发展的概貌，并对其作了探讨性规整，更有益读者合理地选用。

前三章共同组成本书的概论，以期对测量误差分析与不确定度评估及其数据处理所涉及问题有较全面的了解。

第 4 章阐述了测量不确定度评估中概率分析的基本原理，包括客观概率与主观概率两种概念、常见的概率分布及其特征量、概率分布及其特征量的合成、变量测量数据的随机过程描述、常用的随机过程及其分解表示方法等。本章是测量误差分析与不确定度评定所必备的基础性理论。

第 5 章阐述了测量结果及其不确定度评估的数理统计基本方法，如参数估计、区间估计、假设检验、非参数检验等，也即“指南”中涉及的主要统计分析方法。

第 6 章集中阐述了测量结果及其不确定度评估中的现代方法，含一般最小差距方法、稳健统计处理方法、熵分析与熵优化方法、Bayes 统计处理方法、仿真数值分析方法、自助法、支持向量机方法等。也即可对“指南”补充与扩展应用的一些数据处理方法。本章不仅从实用性上论述这些现代方法，还简要地述及各自的理论依据和基本原理以及其所要求的条件和应用场合。

第 7 章论述了在各种不同测量情况下，对常量测量结果的各种相应的最佳评估方法，包括在“指南”中未涉及的对其系统误差的识别与修正、粗大误差的判别与异常数据的剔除及可排除粗大误差影响的测量结果稳健估计方法等。

第 8 和第 9 章两章分别论述了测量不确定度在各种不同测量情况下的各种评估方法，分别包括标准不确定度的 A 类最佳评估与 B 类估算方法，和合成不确定度与扩展不确定度的评估方法等。其中也含有在“指南”中未涉及的或其扩展应用的一些评估方法。

第 10 章简要论述了变量测量中测量结果及其不确定度评估的基本方法，包括其基本理论依据和特点、评定指标、数学模型、分离影响量影响的方案及其数据处理方法等。关于变量动态测量的基本理论及数据处理方法，则不在本书论述范围，可参考作者的另著《动态测试数据处理》。

第 11 章阐述了测量结果及其不确定度表达报告，包括该报告的基本要求、步骤和形式。

第 12 章阐述了各种测量方法的测量结果及其不确定度评估报告的示例，分别结合实例简要述及直接测量、间接测量、测量设备准确度检定、变量测量等测量结果及其不确定度评估的特点与具体评估方法及其报告形式。本章的示例多为作者曾研究或曾参与的项目。

本书得以问世，首先要特别感谢王晓岩女士（现任 Lenovo CIO）的极力推荐与有效地

支持。同时，应感谢中国仪器仪表学会的孟力先生热情地服务与照顾，使得这一已搁置近五年的既有实用价值又有较高学术性的书稿，怕是会像作者曾为出版此书所遇到的种种问题，仍“沉睡”在抽屉里。

还要感谢林春泉、赵玲丽编辑对全书出版更规范化和语气更通俗化所作的修饰。

本书撰写历时较长，不免存在疏漏、谬误，望读者不吝指正。

作者

2008. 11

目 录

前言

第 1 章 概论 1

- 1.1 测量不确定度表述的意义和作用 1
- 1.2 测量误差评估的现状及其问题 2
 - 1.2.1 误差定义及其引起的问题 2
 - 1.2.2 误差分类及其引起的问题 3
 - 1.2.3 误差估算与合成及其引起的问题 4
 - 1.2.4 误差理论发展的现状 5
- 1.3 测量不确定度表述国际标准形成的回顾 6
- 1.4 ISO 1993 (E)“指南”的推广应用 7

第 2 章 基本概念 10

- 2.1 通用测量术语及定义 10
 - 2.1.1 关于量和值 10
 - 2.1.2 关于测量结果及其修正 12
 - 2.1.3 关于准确度 14
- 2.2 测量误差及其类型 16
 - 2.2.1 误差的有关术语及定义 16
 - 2.2.2 误差类型及其定义的变更 17
- 2.3 测量不确定度术语及定义 18
 - 2.3.1 测量不确定度定义及其变更 18
 - 2.3.2 “指南”提出的术语(代号)及定义 19
- 2.4 测量不确定度来源 22
- 2.5 测量不确定度与误差理论的关系 23
- 2.6 对 ISO 1993 (E)“指南”的正确理解及扩展应用 24

第 3 章 测量数据分析与处理 27

- 3.1 被测量分类 27
- 3.2 影响量分类 29
- 3.3 一般测量数据的组成及其类型 30
- 3.4 不同测量数据类型的数学模型 32
 - 3.4.1 拟定测量数据数学模型的基本要求 32
 - 3.4.2 一般测量数据模型 33
 - 3.4.3 被测量类型的模型 36

3.5 测量数据典型示例 36

- 3.5.1 被测常量的测量数据示例 36
- 3.5.2 被测变量的测量数据示例 37
- 3.6 测量数据处理的一般方法 41
 - 3.6.1 测量数据处理的目的和任务 41
 - 3.6.2 选择数据处理方法的基本要求 43
 - 3.6.3 测量数据处理的一般方法和步骤 44
 - 3.6.4 测量数据处理的发展历程 46
 - 3.6.5 测量数据处理方法的归整 48

第 4 章 不确定度评估基本原理 52

- 4.1 不确定度评定中两种概率概念 52
 - 4.1.1 基于频率的客观概率 52
 - 4.1.2 事件关系与概率的性质及运算 56
 - 4.1.3 基于信任度的主观概率 58
- 4.2 不确定度评估的概率分布 59
 - 4.2.1 概率分布规律 59
 - 4.2.2 正态分布 62
 - 4.2.3 常用典型概率分布 64
 - 4.2.4 多元随机变量概率分布和条件概率分布 68
 - 4.2.5 随机变量函数的概率分布 69
- 4.3 概率分布的特征量和特征函数 72
 - 4.3.1 期望——总体均值 73
 - 4.3.2 矩 74
 - 4.3.3 特征函数与累积量 77
- 4.4 概率分布及其特征量合成 79
 - 4.4.1 概率分布及其特征量合成的基本方法 80
 - 4.4.2 确定合成特征量的方法 80
 - 4.4.3 确定合成分布的理论方法 81
 - 4.4.4 确定合成分布的展开法 82
 - 4.4.5 确定合成分布的统示法 83
- 4.5 变量测量数据的随机过程描述 87
 - 4.5.1 变量测量数据的特点 87

4.6 常用随机过程	92	6.2.1 稳健性概念及其基本要求	175
4.6.1 正态随机过程	92	6.2.2 稳健性评价	177
4.6.2 平稳随机过程与遍历性	93	6.2.3 稳健估计量构造方法	180
4.6.3 白噪声	96	6.2.4 稳健估计算法	184
4.7 随机过程的分解表示法	97	6.3 熵分析与熵优化方法	191
4.7.1 随机过程的典型展开式	97	6.3.1 信息熵与随机变量概率分布	191
4.7.2 平稳过程的分解表示法	97	6.3.2 熵的基本性质及运算	196
4.7.3 非平稳过程的平稳化表示法	99	6.3.3 熵与测量不确定度	197
第5章 不确定度评估的数理统计		6.3.4 熵的估计方法	198
方法	101	6.3.5 最大熵方法	201
5.1 数理统计基本概念回顾	101	6.3.6 最小互熵方法与熵优化方法	207
5.1.1 统计推断	101	6.4 Bayes 统计处理方法	212
5.1.2 基本概念	102	6.4.1 Bayes 统计推断的基本原理	212
5.2 参数估计	104	6.4.2 统计决策基本原理	214
5.2.1 参数估计基本方法	104	6.4.3 确定先验分布的方法	220
5.2.2 参数估计的评价	105	6.4.4 Bayes 估计的基本方法	228
5.2.3 参数点估计的常用方法	110	6.4.5 经验 Bayes 估计方法	234
5.3 区间估计	116	6.5 仿真数值分析方法	236
5.3.1 区间估计基本概念	116	6.5.1 仿真数值分析基本原理与方法	236
5.3.2 区间估计方法	117	6.5.2 测量误差概率分布的仿真分析	237
5.3.3 容许区间及其估计方法	119	6.5.3 均匀随机数的生成及检验	239
5.4 假设检验	119	6.5.4 给定概率分布的随机数生成	
5.4.1 假设检验基本概念	120	方法	242
5.4.2 假设检验方法	122	6.5.5 误差概率分布类型的判别方法	249
5.5 非参数检验	126	6.5.6 小样本统计量的概率分布类型	
5.5.1 非参数检验基本概念	127	判别及特征量估计	251
5.5.2 概率分布类型检验——拟合优度		6.5.7 误差分布仿真分析方法的应用	
检验	128	示例	253
5.5.3 正态性检验	132	6.5.8 自助法及其应用	254
5.5.4 基于符号的检验	136	6.6 统计学习理论与支持向量机	268
5.5.5 基于秩的检验	140	6.6.1 概述	268
第6章 测量数据统计处理的现代		6.6.2 统计学习理论的特色	269
方法	150	6.6.3 线性分类的支持向量机	272
6.1 最小二乘法及最小差距法	150	6.6.4 一般分类的支持向量机	278
6.1.1 最小二乘法基本原理	150	6.6.5 回归支持向量机	282
6.1.2 最小二乘法统计特性	154	第7章 常量测量结果的评估	286
6.1.3 最小二乘法的基本算法	155	7.1 评估测量结果的原则及要求	286
6.1.4 最小差距法原理	162	7.1.1 评估测量结果的基本原则	286
6.1.5 最小差距估计的基本算法	164	7.1.2 测量结果最佳估计的要求	287
6.1.6 拟合模型验证	166	7.1.3 分离系统影响和随机影响的基本	
6.1.7 拟合模型的变量选择	171	方法	288
6.2 稳健统计处理方法	175	7.2 “指南”中测量结果的评估方法	289
		7.2.1 “指南”中评估测量结果的前提	

条件	289	9.2.1 扩展不确定度评估方法及其 难点	369
7.2.2 “指南”中测量结果的最佳估计 方法	290	9.2.2 确定包含因子的方法	375
7.2.3 对以均值作为测量结果的评价	291	第 10 章 变量测量结果及其不确定度 的评估	384
7.3 现有测量结果的评估指标	293	10.1 评估变量测量结果及其不确定度的 理论依据	384
7.3.1 常用的测量结果评估指标	293	10.1.1 变量测量不确定度评估的基本 特点	384
7.3.2 测量结果各类有代表性评估指标 的分析比较	295	10.1.2 表述变量测量不确定度的基本 概率统计原理	385
7.4 系统误差的识别与修正	297	10.1.3 评估变量测量不确定度的系统 分析基础	389
7.4.1 系统误差的特点及其对测量结果 的影响	298	10.1.4 随机过程输入系统的响应	394
7.4.2 识别系统误差的方法	301	10.2 变量测量误差分析及不确定度 评定	397
7.4.3 修正系统误差的方法	312	10.2.1 变量测量误差的基本分析 方法	397
7.5 粗大误差判别及异常数据的剔除	312	10.2.2 变量测量误差及不确定度的合理 评定指标	399
7.5.1 粗大误差判别的基本原理	312	10.3 变量测量误差及不确定度评定指标的 估计与检验	401
7.5.2 大样本下粗大误差的判别	314	10.3.1 按集合平均估计法	401
7.5.3 小样本下粗大误差的判别	315	10.3.2 按时间平均估计法	403
7.5.4 剔除异常数据的方法	322	10.3.3 谱密度估计方法	405
第 8 章 标准不确定度评估	325	10.4 变量测量结果与其测量误差的 分离	408
8.1 标准不确定度的 A 类评估方法	325	10.4.1 测量误差对变量测量数据的影 响及其分离方法	408
8.1.1 数据分散性的各种评定方法的 分析比较	325	10.4.2 对变量测量数据分离测量误差 的方案	410
8.1.2 标准差估计方法	328	10.4.3 变量测量数据处理常用 算法	414
8.1.3 测量结果的标准差与置信区间 评估	332	第 11 章 测量结果及其不确定度 报告	416
8.2 标准不确定度的 B 类评估方法	335	11.1 测量结果及其不确定度报告的基本 要求与内容	416
8.2.1 标准不确定度 B 类评定概述	336	11.1.1 一般的要求与内容	416
8.2.2 不确定度 B 类评估的一般方法	337	11.1.2 特定的要求与内容	417
8.2.3 按现有资料的 B 类评估方法	339	11.2 测量结果及其不确定度报告的步骤 与形式	418
8.2.4 按机理分析的 B 类评估方法	342	第 12 章 评估测量不确定度示例	423
8.2.5 按非正态分布处理的不确定度 B 类评估	344		
8.2.6 不确定度 B 类评估的自由度	351		
第 9 章 合成不确定度与扩展不确定 度评估	353		
9.1 合成不确定度评估	353		
9.1.1 不确定度的传递规律	353		
9.1.2 确定灵敏系数的方法	354		
9.1.3 不确定度分量的相关分析与 处理	358		
9.1.4 可略微小不确定度分量	368		
9.2 扩展不确定度评估	369		

12.1	直接测量的不确定度评估示例	423	表 B-8	峰度检验的临界值 γ_{4p} 表	463
12.2	间接测量的不确定度评估示例	429	表 B-9	Sharpiro-Wilk 的 W 检验系数表	464
12.3	测量器具检定的不确定度评估 示例	435	表 B-10	Sharpiro-Wilk 的 W 检验的临界值 W_{α} 表	466
12.4	变量测量的不确定度评估示例	442	表 B-11	D'Agostino D-Y 检验的临界 值表	466
附录 A	符号表	451	表 B-12	样本容量 $n \leq 10$ 的二项分布表	467
附录 B	附表	455	表 B-13	游程数临界值表	468
表 B-1	正态分布 $\phi(z)$ 表	455	表 B-14	Mann-Whitney-Wilcoxon 检验的 U 临界值表	469
表 B-2	t 分布表	456	表 B-15	Spearman 秩相关分析的零相关 检验的临界值表	471
表 B-3	χ^2 分布表	457	参考文献		473
表 B-4	F 分布表	458			
表 B-5	零相关检验的临界值 r_{α} 表	460			
表 B-6	柯氏 D_n 检验的临界值 $D_{n,\alpha}$ 表	461			
表 B-7	偏度检验的临界值 γ_{3p} 表	462			

第 1 章 概 论

自从国际计量局 (BIPM) 及国际标准化组织 (ISO) 等七个国际组织联合公布“测量不确定度指南” ISO 1993 (E)^[1] 以来 (本书简称“指南”, 也多以英语缩写 GUM 简称), 测量结果及其不确定度的评定与最佳估计及其表达方法和规则等已成为国内外在设计、制造、检测、计量确认以及产品质量认证等方面的普遍要求。尽管该指南的贯彻实施已有十几年的时间, 且近年来已进行修订“指南”的多次讨论, 其结果是只做增补而不重新修订。然而, 至今还存在对“指南”有异议或不同见解者。本章仍然首先概述为何要用不确定度表示测量结果的准确度与可靠性, 这样做与以往采用测量误差来表示有何差异, 两者之间存在什么关系, 以及该“指南”的由来与实施及其修订中需要考虑的问题。这将有助于更好地理解与实施“指南”及其修订与扩展应用。

1.1 测量不确定度表述的意义和作用

如前言中简述, 国际计量局 (BIPM) 联合七个国际组织公布的 ISO 1993 (E) “指南”中, 要求国际上统一采用测量不确定度来表征任一被测量的测量结果的准确度, 以便使用者能评价其可靠性, 并在该指南中给出了评定不确定度的具体定量指标和估算的基本方法, 及其表达方法和步骤的通则。同时, 较全面而广泛地论述了这样做的意义和作用。

首先, 由于以往国际上对测量误差及不确定度的表示缺乏一致性, 上至基本物理量基准的准确度表示, 下至通常商业往来中商品质量的表示, 都存在对误差及不确定度定量表示不一致的现象。如有的采用标准差 σ , 有的采用极限误差 Δ_{lim} , 还有些专业领域中采用平均误差、或然误差等等。即便都用极限误差, 有的国家多用 1.96σ (σ 为均方根差) 或 2σ , 即取置信水平 $p \approx 0.95$, 如美国; 有的国家则用 3σ , $p \geq 0.99$, 如过去我国、德国、前苏联等。作者在 20 世纪 70 年代初就曾遇到某印刷厂来咨询其所订购的激光式定位装置要求索赔的实例: 该厂在订购时查看美国某公司产品的对准精度为 $0.2\mu\text{m}$, 而德国某公司的产品则为 $0.1\mu\text{m}$, 于是购得后者, 而开箱后所得说明书中注明该对准精度为 $0.1\mu\text{m}$ ($p = 0.68$)。实则两者的对准精度是同等的, 只不过前者取 2σ 而后者取 σ 而已, 显然无从索赔。此类实例并不罕见。在国际长度基准建立中, 70 年代中期拟选定激光波长作为基准, 而各先进大国所研制的最高准确度的激光器, 其波长不确定度的表示就不一致, 包括对各个不确定度分量的分析和估算, 及其合成为总不确定度等方法各不相同。致使无从比对与评价, 也就无法统一及传递。

上述情况已能说明: 非常需要有一个国际上公认统一的且易于理解、便于执行、又可广泛接受的测量结果的不确定度评定、估算及表述方法。以便可靠地对任一物理量、技术性能参数、科学常数等测量结果的准确度作出一致的定量表征, 才有利于对其质量作比对与评价。ISO 1993 (E) “指南”就是为满足这一需要而诞生的。正如其前言中指出: 此指南的目的是: 促使以足够完整的信息表述不确定度, 为测量结果的国际性比对提供一

个基础。

显然，只有在这样的基础上才能实现准确、可靠地建立各级测量基准和标准；实施计量保证方案（MAP）及执行计量保证方法（MAM）；实现测量设备的计量确认等。这里所指的测量设备是广义的，包括所有测量器具、测量标准、标准物质及其辅助设备，以及进行测量所必须的说明书、规程、规范与所使用的计算机软件等资料。计量确认是为确保测量设备处在满足预期使用要求状态所需的校准、调整、修理、封印、标记等一组操作。这也是当前国内外为保证自由贸易和公平竞争所形成的，受国际贸易牵动的所谓“ISO 9000 现象”的需要（在世界范围内竞相贯彻实施 ISO 9000 系列标准，等价于我国的 GB/T 1900 系列标准）。特别是，该系列标准中，为保证产品、商品质量，均要求其参数准确度符合“指南”所规定的不确定度表示，且与所选配的测量设备准确度的不确定度比为 1/3，最好是 1/10。

可见，测量不确定度是所有产品质量保证体系与测量设备质量保证及剂量确认体系中的重要因素。不确定度的正确评定、合理估算及合成、考虑全部有关数据和资料后详尽而合乎“指南”通则的表述，对提高产品质量与测量准确度水平具有重要作用，也即正确合理地实施该指南，将便于 ISO 9000 系列标准的贯彻实行，及厂矿企业开展质量认证，也有利于我国深化改革开放，发展国际贸易，保证国内外市场接轨，增加经济效益和保护消费者利益等等。

1.2 测量误差评估的现状和问题

如前言所述，在测量结果及其准确度表示上之所以存在不一致现象，主要原因之一在于传统上的测量误差定义与分类引起其评定和估算方法上的一致，并导致测量误差分析及数据处理这一学科发展，至今形成多种观点及不同流派。其中有些见解的分歧属于对误差理论理解深度不一或进展程度各异所造成，无疑对这种情况需求得统一一致，而有些新观点与新方法则属于该学科发展过程中难免出现的正在讨论、尚待完善的问题。显然对后一情况宜“百家争鸣”，不可强求统一。

1.2.1 误差定义及其引起的问题

传统的误差定义很自然地表示为某个量含有误差的值对其无误差的真实值之差。在测试中，测量误差 Δx 定义为被测量的测量结果减去被测量的真值，即

$$\Delta x = \hat{x} - x_0 \quad (1.2.1)$$

式中， x_0 即为测试中待定的未知真实值，且因误差不免存在又是难以确知的。 \hat{x} 为被测量的估计值即测量结果。在一次测量中即为测得值 x_1 ，在多次重复测量中需经数据处理而得，如均值 \bar{x} 或中位值 x_{med} 等。于是测量误差 Δx 也就难以确定，即 Δx 针对测量结果来表示而 $\hat{x} = ?$ ，又因 x_0 不可确知，即 $x_0 = ?$ 。其结果显然是测量误差的表示形式及大小均难以确定， $\Delta x = ?$ 。因而存在如何评定测量误差的问题。

测量误差是表征测量结果准确度、评定测量方法和测量器具准确度的定量指标。为了表示测量结果的可信度及测量方法和测量器具的可靠性，就必须客观而科学地分析与定量估算

出测量误差。于是出现不同主张，见下例。

例 1.2.1 在测长仪上 8 次重复测量 $\Phi 20 \pm 0.005\text{mm}$ 的滑块宽度，其数据 x_i (mm): 19.999、20.005、20.000、19.999、19.998、19.999、19.998、20.000。

方案一：最常用者，测量结果取均值 \bar{x} ，测量误差取以 3 倍标准差 s 表示的极限误差 Δ_{lim} ，即

$$\begin{aligned}\hat{x} &= \bar{x} = n^{-1} \sum_{i=1}^n x_i = 20.00075\text{mm} \\ s &= [(n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2]^{1/2} = 2.1\mu\text{m} \\ \Delta_{\text{lim}} &= 3s = 6.3\mu\text{m}\end{aligned}$$

方案二：考虑 $x_2 = 20.005\text{mm}$ 偏大。可能为异常数据，而取中位数 x_{med} 为测量结果，相应地取中位绝对差 MAD 表示测量误差，即

$$\begin{aligned}\hat{x} &= x_{\text{med}} = \text{Med}_i[x_i] = 19.999\text{mm} \\ \text{MAD} &= \text{Med}_i|x_i - x_{\text{med}}| = 1\mu\text{m}\end{aligned}$$

式中，Med 为取中位数运算，简称中位算子。

方案三：测量结果取众数 x_{mod} ，即出现频数最多者，而测量误差取最大差，即

$$\begin{aligned}\hat{x} &= x_{\text{mod}} = 19.999\text{mm} \\ e_{\text{max}} &= \max_i|x_i - x_{\text{mod}}| = 6\mu\text{m}\end{aligned}$$

还有其他方案，不再罗列。上述方案中诸代号及其算法详见后述。

可见，不同评定测量结果与测量误差的方案所得的数值各异，其中测量误差数值差异较大，很有必要统一用不确定度评定。然而，其中个别数据偏大或偏小是否正常，由剔除与否而引起的差异，则不应视为评定不统一，而属于合理数据处理的问题。

1.2.2 误差分类及其引起的问题

误差来源不同则对测量结果的影响就各异，因此需要对误差进行分类处理。传统上多按误差影响性质的不同而分为两种基本类型，即系统误差和随机误差。有时还将偶尔产生的某种不正常原因所造成数值特别大的误差，称为粗大误差，其所影响的数据称为异常数据。

系统误差是指在测量过程中以恒定数值及正负号或以确定性规律影响其测得数据的误差，有时也称其为规律误差。对系统误差的处理是力求确定其值（包括数值和正负号）或其变化规律，再对测量结果或测量数据予以修正。

随机误差是在测量过程中以不规则的数值及正负号随机性地影响其测得数据的误差，也称之为偶然误差。然而，这种误差并非毫无规律，只不过在每次测量或测量过程中，每一瞬间所呈现的不规则或其数值及正负号不可预知，而就其整体而言却呈现出某种统计规律。对随机误差的处理只能是评估其对测量结果影响的数值范围，力求削减其影响而难以修正。

对于粗大误差，因其对测量结果有所歪曲，常通过判别而剔除其所影响的异常数据。

上述的误差分类及其含义是早期作出的，并非当前作出的定义。作者认为虽然未对其作出定量的定义，可是概念却是明确的，即着重从影响性质上划分，只不过需补充其可操作性

而已。

随着科技的发展，对测量准确度的要求不断提高，从而对误差评估及数据处理也要求更加严密。于是在深入研究系统误差的计算及修正和对随机误差的评估中，出现了因理解上的差异，而形成分析及处理上各辟蹊径、自成流派的现象。

首先对系统误差按其影响规律的不同，再分为常量系统误差和变量系统误差（也有分别称之为定值和变值者），甚至后者再细分为线性变化、周期变化等等。显然这样细分作者认为无可非议，因为不致引起混淆。

然而，这些影响规律事先未必已知，且同一误差源在不同的测量方式和不同的测量条件下，其影响规律也未必相同，又因情况复杂以致难以确定也不易求得，于是又出现已定系统误差和未定系统误差的划分。对未定系统误差，随着处理方法的不同，又有随机性系统误差^[7,8]、双向系统误差、系统性随机误差或前次随机误差^[9]、半系统误差^[10]、不定常差^[11]等等称谓，并曾对国内误差分析界有过一定的影响。

所谓随机性系统误差是指未知的或虽可知却过于复杂的系统误差，而按随机误差来处理。简言之，未定系统误差源按随机误差处理。所谓系统性随机误差则指前次随机误差源在本次测量中为未定系统误差者，也只得按随机误差处理。所谓半系统误差定指仅知误差范围的未定系统误差，也是按随机误差处理。所谓不定常差则指仅知常量数值而未知其正负号的系统误差，而按两点分布随机误差处理者。在对这些误差举例中，有标尺刻线误差、分度误差等。总之，不论这些误差分类如何称谓，均归结为具有随机性即不确定性而按随机误差处理。

对于随机误差，除正常地划分为：正态和非正态、对称和非对称等分布类型外，也存在纠缠系统误差源在本次为随机误差之类的问题。

对上述情况，“指南”中统以具有概率分布意义的不确定度来评定，而免作误差分类及不同分类的不一致处理方法之类的纠缠，以避免有主观随意性或理解不一之弊。

1.2.3 误差估算与合成及其引起的问题

由误差定义及分类引起的问题显然会延伸至误差估算与合成中。例 1.2.1 中已反映出对误差评定及估算所造成的差异。这种差异主要因误差定义仅有理论意义，或只是概念化的定义而未定量化，致使在误差评定及估算上有不同见解及不同方法。再因误差源多而复杂，并非现有数据处理方法都能将误差分离出和估算出。况且还有些误差源的影响并未含于现有的或本次测量数据中，只能另用其他方法估算，包括用已有的经验信息和资料去推测。在误差分类中出现未定系统误差或随机性系统误差、系统性随机误差等等多属此类。

此外，对误差理论及数据处理方法掌握程度的不同显然也会导致误差估算及合成的差异。如掌握不足者，其评定指标偏简易而估算方法偏“保险或安全”，像例 1.2.1 中方案三采用 e_{\max} 那样；掌握较深者，其评定指标科学而估算方法合理，像例 1.2.1 中方案一采用通过 s 估算出与 e_{\max} 相当的 Δ_{lim} 那样。而 e_{\max} 比 s 的统计量估计准确性及效率差些，后者通常为最佳估算。这正是不确定度以标准差为评定指标的原因之一。

在误差估算及合成中可归结为三种情况：对确定性的量多用解析法；对随机性的量多用统计法；对未知的量则不论其属性为何种，均只能用其他方法去确定，包括利用现有的一切经验信息来推测。这种划分主要针对数据处理而言，利用测试技术措施测定误差的方法则不

在本书的论述范畴内。

对表现为确定性规律的系统误差除测定方法外，主要靠解析法确定，其合成均采用代数法和，如对已定系统误差等修正及合成即此。

对表现为统计规律的随机误差，则靠概率统计方法去估算及合成。由于方差 σ^2 表征随机变量的离散度，故绝大多数人认为以标准差 σ 或 s 即 σ 的无偏估计量作为随机误差等评定指标较合理，其合成则按方差和或标准差平方和规律（存在相关时应考虑协方差项）。即使采用其他统计量为评定指标，也多按该规律去换算。这是当前认识较为一致的处理方法。

问题和分歧均出自上述第三种对未知量的处理上，包括未定系统误差和未影响到本次测得数据随机变动的随机误差，即前述的所谓系统性随机误差、随机性系统误差、半系统误差、不定常差等等。对这类误差多从保险观点采用最大误差或误差限为评定指标，像 e_{\max} 或类似 Δ_{lim} 那样，却不是统计法估算出，而是以其他方法按经验资料推测所得。在合成方法上则出现的差异更多：有主张与随机误差同样合成的；有的则主张在合成中其分布取均匀分布；也有主张单独用绝对值和法合成后，再与随机误差合成或相加等等。如在参考文献[10]中经细致分析后竟归纳总结出三类11个合成公式，且建议了其选用原则。在参考文献[8, 9, 11, 12]中也都推荐了各自认为适宜的合成方法和公式。显然这些不同见解必然导致误差估算和合成的结果呈现差异，以致不利于比较与评价。因此，“指南”中给出了统一的不确定度评估和合成的基本方法。

1.2.4 误差理论发展的现状

尽管误差分析及数据处理存在上述各种问题，误差理论自从高斯（Gauss, C. F.）于19世纪末创建以来，毕竟经历了200余年的发展，已形成较完整而成熟的理论体系。且随着科技的不断发展及准确度要求的不断提高，误差分析及数据处理仍处在不断发展和完善的过程中，目前又提出了许多新观点和新方法。如：

1) 从传统的正态分布误差的深入研究与统计处理扩展至非正态分布误差尤其是非对称分布误差的分析与统计处理。因为实践表明，非正态误差是客观存在的，且并非少见^[13]。因而近年来已广泛地开展对非正态误差的各种概率分布形式的描述、评定指标、估算及合成等各方面的探讨，尤其对利用高阶矩或累积量（cumulant）的分析方法与概率分布的级数展开法和统一表示法的研究及应用^[3, 6, 7, 9, 12-15]。

2) 在统计处理方法上，在传统的最小二乘法基础上，不仅为了适应各种不同应用场合，而改善和发展其处理方法。如采用正交变换、特征值或奇异值分解等算法，各种形式的加权处理等。且为适应各不同专业领域的技术要求、不同数据统计特性的条件等，而扩展至其他最小范数或最小距离准则下的处理方法。如残差绝对值和为最小的最小一乘法、残差最大值为最小的最小 α 乘法、或该两者之间的最小 m 乘法，以及其他的最小距离法等。还有非线性模型的处理方法等^[11, 16, 17]。

3) 从传统的最佳估计或最佳拟合的统计处理扩展至稳健统计（robust statistic）处理。因为实际数据常难满足独立性、正态性、无异常外部干扰（或称“污染”）等最佳性假定条件，而偏离这些条件，所采用的最佳估计或拟合方法将失去其最佳性，甚至会失效。而稳健估计和拟合方法可在数据稍有偏离原假定的概率分布模式，及受少量粗大误差或一些异常的小误差污染下，仅使其估计或拟合结果作较小的改变，仍基本上保持原有的最佳性而不致失

效。故稳健统计已成为现代数据处理中颇具活力的分支^[13,16,18~23]

4) 从传统上只依靠现有数据作统计处理扩展至与现有数据一起再充分利用已有的知识、经验、资料等先验信息, 而采用 Bayes 统计处理方法。尤其在小样本下, 只要先验分布选定得当或采用经验 Bayes 方法等, 可使 Bayes 方法结果更为准确、可靠。不确定度的 B 类评定方法就考虑到应用这一统计原理^[1,24,25]。

5) 从传统的统计分析方法扩展到信息熵分析及熵优化准则的方法。依据熵可作为信息不确定的惟一性度量, 熵最大就意味着最大不确定性, 以及每种随机变量概率分布都对应着一个熵值(逆转对应并不成立)等原理, 即可用误差熵值反映其离散度, 形成熵分析方法。又为避免求解具有多种可能解的各种不适定问题中作主观性假定, 即只依据所得的数据含有的全部信息, 及解题所要求的约束条件, 此外不再作任何主观假定, 在最不确定性即最大熵准则下求出不适定问题的解, 简称最大熵方法。进而, 在依据数据及所要求的约束条件上, 又有已知的知识、经验、资料等先验信息可用时, 则可使两种信息最大限度地相符合, 及相互间的不确定性应最小, 也即按最小互熵准则来解题, 简称最小互熵方法。并可统归为熵分析与熵优化方法。这种方法的主要优点在于可不涉及概率分布及其他主观假定。目前已在各种专业领域中应用^[26~29]。

6) 从传统的静态测量数据处理扩展到动态测量数据处理。传统上基于随机变量统计方法的静态测量数据处理在变量的动态测量广泛应用下已不适用, 已扩展至基于随机过程的统计方法。尤其是长过程测量, 包括变量测量过程和常量重复测量过程(如在线测量或质量控制中的长期监测等), 为适应其未知的复杂变化规律, 及跟踪分析、处理和显示其时变统计特性(即特性量、技术参数或评定指标等), 出现了各种自适应统计处理方法, 包括各种递推式算法(recursive algorithm), 以及近年来兴起的着重于精确描述上的移动式算法(moving algorithm)。

还需特别指出, 在计算机及其各种算法软件广泛应用与普及后, 用于分析解决各种难题的计算机数值模拟或仿真分析方法, 尤其是仿真输出结果的分析技术, 在误差分析及数据处理中已常用于开发新分析技术和数据处理方法及其效能的验证等。这就形成了从传统上应用概率统计方法而扩展为非统计方法, 如上述基于信息量分析的熵分析和熵优化方法; 基于模糊集和模糊逻辑的模糊分析方法; 具有多分辨和变尺度的小波分析方法; 模拟生物进化、遗传等的遗传进化算法; 模拟生物神经思维的神经网络分析方法等等。

上述的一些新方法多在 20 世纪 70 或 80 年代以后发展至实用化的, 进而统称为现代方法。显然, 在应用现代方法中不免会出现不一致问题, 应视为发展中的客观存在, 不宜强求统一。在“指南”中也提及和应用了上述的某些现代方法, 还有待扩展及补充之处。

1.3 测量不确定度表述国际标准形成的回顾

鉴于测量误差评估中存在各种不一致现象, 已影响到对测量结果及其准确度作国际上的比对与评价, 尤其已影响到基本物理量国际基准的建立。1978 年, BIPM 应世界最高计量权威——国际计量委员会(CIPM)的要求, 向 32 个国家的有关计量部门和实验室发