

国家计量技术法规统一宣贯教材

测量不确定度理解 评定与应用

国家质量监督检验检疫总局计量司 审定

叶德培 编著

JJF 1059—1999

JJF 1059—1999

JJF 1059—1999

中国计量出版社

CHINA METROLOGY PUBLISHING HOUSE



国家计量技术法规统一宣贯教材

测量不确定度 理解 评定与应用

国家质量监督检验检疫总局计量司 审定

叶德培 编著

中国计量出版社

图书在版编目(CIP)数据

测量不确定度理解 评定与应用/叶德培编著. —北京:中国计量出版社, 2007.9

国家计量技术法规统一宣贯教材

ISBN 978-7-5026-2714-0

I. 测… II. 叶… III. 测量-不确定度-教材 IV. TB9

中国版本图书馆 CIP 数据核字 (2007) 第 129167 号

内 容 提 要

本教材是原国家质量技术监督局计量司组编的国家计量技术法规统一宣贯教材《测量不确定度评定与表示指南》(中国计量出版社 2000 年 4 月出版)的补充,总结近十年来我国在测量不确定度应用方面的实践经验和问题,尤其是探讨了测量不确定度在合格评定中的应用,量值比对时测量不确定度的考虑,科研项目方案论证时不确定度的预估,测量不确定度在测量能力表达时的应用,测量仪器的不确定度及计量标准的不确定度问题以及证书报告中测量不确定度的说明等问题;根据 ISO 出版的《测量不确定度表示指南》(GUM)进一步加深对测量不确定度的理解及应用,旨在更广泛地宣传和贯彻国家计量技术规范 JJF1059—1999。

为方便读者学习使用,本书配有作者讲课的音像资料。

本书可供科研、生产和教学等有关人员使用,可以作为注册计量师的学习材料,是从事计量、标准、质量及实验等人员的必备资料。

中国计量出版社出版

北京和平里西街甲 2 号

邮政编码 100013

电话 (010)64275360

<http://www.zgjl.com.cn>

北京市黎明印刷厂印刷

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 12.5 字数 303 千字

2007 年 10 月第 1 版 2007 年 10 月第 1 次印刷

*

印数 1—2 000 定价:88.00 元

前 言

自国家计量技术规范 JJF1059—1999《测量不确定度评定与表示》于 1999 年颁布以来,测量不确定度在我国计量界已得到广泛应用,并且逐渐波及我国各科学技术领域和测量界。鉴于产品质量、司法裁定、科研成果、技术水平、国际贸易等方面越来越需要用测量数据来说话,人们比以往任何时候都更加关注测量结果的可信程度。改革开放促进了我国与国际的广泛交流和商品贸易,必然要求我们要采用与国际接轨的测量不确定度来表示测量结果的质量。

本教材是 2000 年 4 月出版的国家计量技术法规统一宣贯教材《测量不确定度评定与表示指南》的补充,作者总结了近十年来我国在测量不确定度应用方面的实践经验和问题,尤其是探讨了测量不确定度在合格评定中的应用,量值比对时测量不确定度的考虑,科研项目方案论证时不确定度的预估,测量不确定度在测量能力表达时的应用,测量仪器的不确定度和计量标准的不确定度问题以及证书报告中测量不确定度的说明等问题;并根据 ISO 出版的《测量不确定度表示指南》(GUM),进一步加深了对测量不确定度的理解及应用,旨在更广泛地宣传和贯彻国家计量技术规范 JJF1059—1999。

本教材内容主要包括:规范的适用范围、基本术语及其概念、测量不确定度的评定方法、测量结果及其测量不确定度的报告形式、测量不确定度的应用介绍以及测量不确定度评定的举例。最后还给出了两个附录:附录 1 是测量不确定度评定中常用的公式;附录 2 是对 ISO 出版的《测量不确定度表示指南》(GUM)原文(包括 GUM 中的全部附录)的译文。

为方便读者学习使用,本书还附有作者讲课的音像资料。

本书可供科研、生产和教学等有关人员使用,可以作为注册计量师的学习材料,是从事计量、标准、质量及实验等人员的参考资料。

为了进一步搞好对测量不确定度的宣贯,对于书中不足和错误之处,恳请全国广大读者提出宝贵意见,并将意见寄送或转告国家质量监督检验检疫总局计量司。

国家质量监督检验检疫总局计量司

宣 湘

2007 年 8 月

目 录

第一章 概 述	(1)
第一节 引 言	(1)
第二节 评定测量不确定度的意义	(2)
第三节 测量不确定度的发展历史	(3)
第四节 测量不确定度的适用范围	(5)
第二章 基本术语及其概念	(7)
第一节 基本的概率与数理统计术语及其概念	(7)
第二节 基本计量学通用术语	(21)
第三章 测量不确定度的评定	(30)
第一节 测量不确定度评定的一般要求	(30)
第二节 建立测量的数学模型	(31)
第三节 分析测量不确定度的来源	(33)
第四节 评定标准不确定度分量	(36)
第五节 计算合成标准不确定度	(43)
第六节 确定扩展不确定度	(52)
第四章 测量结果及其测量不确定度的报告形式	(56)
第一节 测量结果及其不确定度的报告	(56)
第二节 合成标准不确定度的报告形式	(57)
第三节 扩展不确定度的报告形式	(57)
第四节 测量结果及其不确定度的有效位数	(58)
第五节 符号的正确书写	(61)
第五章 测量不确定度应用的介绍	(62)
第一节 中国合格评定国家认可委员会关于应用测量不确定度的要求	(62)
第二节 测量不确定度在合格评定中的应用	(63)
第三节 实验室间量值比对时测量不确定度的考虑	(67)
第四节 科研项目方案论证时不确定度的预估	(75)
第五节 测量不确定度在测量能力表达时的应用	(75)
第六节 测量仪器的不确定度及计量标准的不确定度	(76)
第七节 证书报告中测量不确定度的说明	(77)
第六章 测量不确定度评定举例	(79)
第一节 圆柱形工件直径的测量	(79)
第二节 容器内温度的测量	(80)

第三节	圆柱体体积测量	(83)
第四节	微波衰减器的衰减量校准	(85)
第五节	脉冲产生器通道时延校准	(86)
第六节	量块的校准	(88)
第七节	样品中所含氢氧化钾的质量分数测定	(92)
第八节	数字多用表的检定	(94)
第九节	确定电阻的温度修正值	(97)
第十节	高稳晶振日漂移率的测量	(99)
附录 1	不确定度评定中常用的公式	(103)
附录 2	测量不确定度表示导则	(105)
参考文献		(194)

第一章 概述

第一节 引言

本书以 GUM1995 版为基础对测量不确定度的理解、评定、表示与应用作了介绍,以便在更广范围内宣传和贯彻国家计量技术规范 JJF1059—1999《测量不确定度评定与表示指南》,旨在有助于我国注册计量师、计量科研人员及除计量界外需要了解测量不确定度的人员了解和学习测量不确定度。

贯彻关于测量不确定度评定与表示规范的目的:

1. 在我国推广使用国际上已经广泛使用的测量不确定度

《测量不确定度表示导则》(Guide to the Expression Uncertainty in Measurement 简称 GUM)是由国际标准化组织(ISO)等七个国际权威组织联合发布的,自 1993 年以来,经 10 多年的推广和应用,现已在国际上广泛使用,成为各国在表示测量结果时统一遵循的准则。国家计量技术规范 JJF1059—1999《测量不确定度评定与表示》于 1999 年颁布,该规范的制定也为测量不确定度在我国的应用起到了推动的作用。但在许多领域还刚刚起步,仍需要进一步澄清模糊概念,正确理解和使用测量不确定度。

2. 以科学合理和完整的信息给出测量结果

当完成测量时,应该给出测量结果。如果给出测量结果时未给出其可信程度或可信的范围,这种测量结果是不完整的。因为测量结果是否有用,很大程度上取决于其可信程度,它是测量结果质量的描述。如何给出完整的测量结果?以前,我们用测量误差来表示,由于测量误差是测量结果与真值之差,真值往往是未知的,这种表示方法虽然我们长期使用过,但现在看来是不够科学的。现在国际上约定的做法是用测量不确定度来表示。测量结果必须有不确定度说明时,才是完整的和有意义的。

3. 规范测量不确定度的评定与表示方法

国家计量技术规范 JJF1059 在评定和表示测量不确定度的方法,包括所用的符号等方面,等同采用 GUM 规定的方法,以便与国际接轨。GUM 的方法有以下优点:

(1)内部协调一致性:测量结果的不确定度可以从对不确定度有贡献的分量导出,与这

些分量怎样分组没有关系，也与这些分量如何进一步分解为下一级分量无关。

(2)可传递性：当上一个测量结果用于下一个测量时，其不确定度可作为下一个测量结果的不确定度分量。

(3)具有广泛的适用性：在诸如工业、商业及与健康或安全有关的一些领域中，往往要求提供较高概率的置信区间，该方法能方便地给出这样的区间及相应的概率。

第二节 评定测量不确定度的意义

一、测量的重要性

测量是在科学技术研究、工农业生产、国内外贸易、工程项目以及日常生活的各个领域不可缺少的一项工作。人们每天都在进行大量的测量，在测量司法执法、维护消费者权益、保护资源和环境、医疗卫生、国防建设等方面测量起着越来越大的作用。随着当前经济的全球化趋势和我国经济的飞速发展，测量结果是否可信，或者说测量的质量如何，是人们极其关心的问题。测量的准确性可能直接影响到国家和企业的经济利益。如果我们的出口货物由于称量不准，有可能引起经济纠纷并导致赔偿，造成很大的损失。测量的结果是科学研究成果的评价依据，也是产品检验合格判定、司法裁定等裁判的依据。凭什么说某项成果达到了国际先进水平？凭什么说该项产品不合格？凭测量结果才是最有说服力的。测量的质量还往往成为科学试验成败的重要因素。测量的质量也会影响到人民的健康和安全。在使用激光治疗仪治疗疾病时，若对剂量测量不准，剂量太小，达不到治病的目的，延误治疗；剂量过大会造成对人体的伤害。由测量结果得出的结论还可能成为决策的重要依据。例如，决策投资方向、决策是否可以作发射试验、决策采取什么样的应对措施等。错误的或不可信的测量结果会动摇决策者的决心，延误时机。

二、规范测量不确定度评定和表示方法的必要性

(1)完整的测量结果应包括测量不确定度的说明，使人们能够了解该测量结果的可信程度。在市场竞争激烈、经济全球化的今天，测量不确定度评定与表示方法的统一，乃是科技交流和国际贸易的迫切要求。

(2)我国用统一的准则对测量结果及其质量进行评定和表示是与国际接轨的需要，也是我国经济发展的必然趋势。它有利于：

- ①测量结果间的比较；
- ②科学技术成果的评价与交流；
- ③商品贸易中减少技术壁垒和避免不必要的误会；
- ④对计量标准、标准物质和标准参考数据的评定与发布；
- ⑤用户对计量技术机构校准或检测服务后给出的证书或报告的理解；
- ⑥校准或检测实验室的技术能力认可和国际互认；
- ⑦生产中的质量控制以及进行质量体系认证时对产品质量保证能力的评价；
- ⑧有利于根据测量结果做出有效的决策等。

第三节 测量不确定度的发展历史

1. 测量不确定度的发展

早在 1963 年美国国家标准局(NBS)的数理统计专家埃森哈特(Eisenhart)在研究“仪器校准系统的精密度和准确度的估计”时提出了定量表示不确定度的概念和建议,受到了国际上的普遍关注。

20 世纪 70 年代, NBS 在研究和推广测量保证方案(MAP)时在不确定度的定量表示方面有了进一步的发展。不确定度这个术语逐渐在测量领域广泛使用,但具体表示方法方面很不统一,并且不确定度与误差同时并用。

1977 年 5 月国际电离辐射咨询委员会(CCEMRI)的 X- γ 射线和电子组讨论了关于校准证书如何表达不确定度的几种不同建议,但未作出决议。1977 年 7 月的 CCEMRI 会上提出了这个问题的迫切性, CCEMRI 主席——美国 NBS 局长 Amber 同意将此问题列入送交国际计量局的报告,并且,由他作为国际计量委员会(CIPM)的成员向 CIPM 发起了解决测量不确定度表示方面的国际统一问题的提案。

1977 年, CIPM 要求国际计量局(BIPM)联合各国国家计量院着手解决这个问题。1978 年 BIPM 就此问题制定了一份调查表,分发到 32 个国家计量院及 5 个国际组织征求意见。1979 年底得到了 21 个国家实验室的复函。

1980 年, BIPM 召集和成立了不确定度表述工作组,在征求各国意见的基础上起草了一份建议书: INC-1(1980)。该建议书向各国推荐了测量不确定度的表述原则。自此,得到了国际初步统一的测量不确定度的表示方法。

1981 年,第 70 届国际计量委员会批准了上述建议,并发布了一份 CIPM 建议书: CI-1981。1986 年, CIPM 再次重申采用上述测量不确定度表示的统一方法,并又发布了一份 CIPM 建议书: CI-1986。CIPM 建议书推荐的方法是以 INC-1(1980)为基础的。CIPM 要求所有参加 CIPM 及其咨询委员会赞助下的国际比对及其他工作中,各参加者在给出测量结果的同时必须给出合成不确定度。

20 世纪 80 年代以后, CIPM 建议的不确定度表示方法首先在世界各国的计量实验室中得到广泛应用。但正如国际单位制计量单位不仅在计量部门使用一样,测量不确定度应该可以应用于一切使用测量结果的领域。如何进一步推广使用的问题提到了日程上。

1986 年 CIPM 要求国际标准化组织(ISO)能在 INC-1(1980)建议书的基础上起草一份能广泛应用的指导性文件,该项工作得到了 7 个国际组织的支持,该 7 个国际组织是: ISO(国际标准化组织)、IEC(国际电工委员会)、CIPM(国际计量委员会)、OIML(国际法制计量组织)、IFCC(国际临床化学联合会)、IUPAC(国际理论化学和应用化学联合会)、IUPAP(国际理论物理和应用物理联合会)。这 7 个国际组织包括两个权威的标准化组织、两个权威的计量组织和三个物理、化学、医学测量方面的权威组织。自此,成立了专门的工作组即国际标准化组织(ISO)的第四技术顾问组(TAG4)第三工作组(WG3),开始起草“测量不确定度表示导则”,该工作组的成员是由 BIPM、ISO、IEC 和 OIML 四个国际组织提名的。

1993 年,经过工作组近 7 年的努力,完成了“测量不确定度表示导则”的第一版,并

以 7 个国际组织的名义联合发布, 由 ISO 正式出版发行。同时终止了 ISO/TG69/SC6/WG3 关于测量不确定度标准的起草工作。

1995 年在对“测量不确定度表示导则—1993e”作了一些更正后重新印刷, 即《Guide to the Expression of Uncertainty in Measurement—corrected and reprinted, 1995》(简称 GUM), 为在全世界采用统一的测量结果的不确定度评定和表示方法奠定了基础。该指导性文件已经使用了 10 多年, 目前为止它仍然是有效版本。

2. GUM 的特点

(1) 因为 GUM 是在由 7 个权威组织历经 7 年反复研究讨论和征求各国意见的基础上制定的, 因此 GUM 具有国际权威性。

(2) GUM 是在 INC-1(1980)、CI-1981 和 CI-1986 基础上编制的应用指南。因为是指导性技术文件, 它可以免去作为“标准”编写时的许多文字和内容上的局限性, 在术语定义、概念、评定方法和报告时的表达方式上都作了进一步的统一规定, 并有许多解释性的内容。利用附录的形式还回答了许多应用时所遇到的问题, 并给出了许多应用举例。具有很强的操作性和实用性。

(3) GUM 代表了当前国际上在表示测量结果及其测量不确定度时的约定做法。使全世界不同国家、不同地区、不同学科、工程、商业、工业、法规等领域在表述测量结果和测量不确定度时具有一致的含义, 便于理解、翻译和比对, 它对推动科技进步和促进国际交流具有重要意义。

(4) 现在, 各国都将 GUM 方法转化为本国标准或技术规范加以推广应用。为正确执行 GUM 方法, 许多实验室或计量组织, 例如美国的 NIST, 制定了本单位的实施指南。一些区域性和全球性的国际组织, 例如亚太地区计量组织 (APMP)、欧盟计量组织 (EUROMET)、国际实验室认可组织 (ILAC)、亚太实验室认可组织 (APLAC) 及欧盟认可合作组织 (EA) 等, 也都强调用 GUM 方法来表示测量结果及其测量不确定度。在国际杂志上发表的论文或评论以及校准证书和测试报告等文件上, 基本上已都采用了测量不确定度。测量不确定度已经被越来越多的人所理解和应用。

1999 年 1 月我国发布的国家计量技术规范 JJF1059—1999《测量不确定度评定与表示》就是在这样一个背景下制定的, 它以法规形式规定了我国贯彻 GUM 方法的具体要求。以便在测量结果及其不确定度的评定与表示方法上与国际接轨, 以利于我国的国际交往和经济发展。自规范发布以来, 测量不确定度在我国已经得到广泛应用, 许多同志为宣传和贯彻测量不确定度做出了很多努力, 许多单位与个人在总结应用经验的基础上写了不少有价值的文章和书籍。

测量不确定度的概念以及不确定度的评定和表示方法的采用, 是计量科学的一个新进展。从 1963 年提出测量不确定度概念, 到 1993 年正式发布能统一全世界关于测量不确定度应用的指导性文件 GUM, 整整花费了 30 年时间, 它汇集了世界各国计量学家的经验和智慧, 并在与守旧的习惯力量做不懈争辩的过程中诞生。经过多次修改、完善和改进, 最后形成了目前这样较为科学、系统和完整的指导性文件。但是与所有的事物一样, 随着社会的进步, GUM 也会有进一步的发展。人们在总结应用中的经验和问题后, 对 GUM 作进一步修订是必然的。我们每一个人都可以积累这方面的经验和意见, 把我国的想法收集起来, 积极参与到 GUM 的修订工作中去, 使 GUM 更实用, 也为世界做出我们应有的贡献。

第四节 测量不确定度的适用范围

一、适用的领域

测量不确定度评定与表示的通用规则适用于各种准确度要求的各类测量领域。因而也将渗透到科学技术的各个领域。包括：

(1) 建立国家计量基准和各级计量标准

适用于在建立计量基准或各级计量标准时，评定和给出其复现的标准量值的测量不确定度。

(2) 计量标准装置间的国内外比对以及检测设备的实验室间比对

适用于计量标准装置间或检测设备间的比对，参与比对的各方在给出测量结果时必须按照统一的要求同时给出测量不确定度。通过对参加比对的各实验室所得数据的处理，可以得出测量结果一致性的评价。比对结果是测量结果可信度的证明，也是对实验室技术能力的一种验证。

(3) 标准物质的定值，标准参考数据的发布

适用于标准物质按规定的方法定值后，其标准值连同其不确定度的发布。也适用于需要说明不确定度的标准参考数据的发布。

(4) 测量方法、检定规程、校准规范等技术文件或标准的编制

编制测量方法、校准规范和检定规程时，应该分析和评定该方法的测量不确定度，以便使用者在分析测量结果的不确定度时作为参考或作为一个分量加以使用。

校准等级关系图又称测量的溯源等级图，是说明从工作测量器具向上溯源到各级计量标准直至国家基准的溯源链的不确定度关系图的技术文件，图中需明确标明溯源链中各级的测量不确定度，并符合有关的比例关系要求。

(5) 科学技术研究及工程领域的测量

测量不确定度适用于一切科技与工程项目，这是一个非常广阔的应用领域，例如：无论是科学发明还是技术创新，所有的科技成果往往都必须以测量结果及其不确定度来评价其水平。重大工程的方案论证离不开测量不确定度的分析和预估，从而给出合理的技术要求。工程的验收大纲应该规定测量的要求包括测量不确定度的要求。高等学校学生在毕业论文涉及到测量结果时也应该正确使用测量不确定度，因此关于测量不确定度的知识也适用于大专院校的测量课程。

(6) 计量认证、计量确认、质量认证以及实验室认可

在计量认证、计量确认、质量认证中，要根据相关的标准，对测量设备能否满足产品质量检测的要求、测量结果及其不确定度能否满足使用的要求进行评审；在实验室认可中，对测量范围及测量不确定度的考核是对该组织的技术能力的评定。

(7) 测量仪器的校准和检定

测量仪器是人们测量时必不可少的工具，为了保证其质量满足使用要求，必须进行定期校准或检定。也就是将测量仪器与相应的计量标准进行技术比较，从而给出仪器的校准值、校准曲线或修正值、修正曲线，此时应该同时给出这些值的测量不确定度；对于法制计量范

围内的测量仪器必须按规定与相应的计量标准进行技术比较后,再与被检测量器具的技术指标作比较,给出合格或不合格的检定结论,此时应该在检定证书中给出标准值或所用计量标准的测量不确定度,因为它关系到合格评定的可信程度或误判风险。

(8)产品或商品的检验和测试

所有的产品或商品都需要经过检验,合格后才能投入市场。凡是需要测量,包括检验、检疫中需要定量测量的部分,须记录测量结果和测量不确定度,以备质量追溯。虽然用户并不都需要知道如何测量及测量不确定度有多大,只相信是否合格的结论,但是到发现具体质量问题时,甚至需要处理投诉意见或法庭裁定时,记录的完整性对于分析问题和解决问题是很重要的。

(9)生产过程的质量保证

适用于制造厂在生产过程中进行的质量控制和产品检验时对测量结果的表述及合格判定。

(10)贸易结算、医疗卫生、安全防护、环境监测及资源测量

也适用于各项强制检定项目中对计量标准及检测设备的要求及对测量结果的质量的评定。

二、国家计量技术规范 JJF1059—1999 的适用范围

(1)规范适用于涉及有明确定义的,并可以用惟一值表征的被测量估计值的不确定度。

例如:直接用数字电压表测量频率为 50Hz 的某实验室的电源电压,电压是被测量,它有明确的定义和特定的测量条件,用的测量仪器是数字电压表,进行 3 次测量,取其平均值为测量结果,测量结果为 220.5V,它是被测量的估计值并用一个值表征的。现有规范对这样的测量结果进行测量不确定度评定和表示是适用的。

又如:通过对电路中的电流和电压的测量,用公式计算出功率的测量结果的间接测量,也符合有明确定义的并可用惟一值表征的条件,因此本规范是适用的。

(2)当被测量为导出量,其数学模型即函数关系式中的多个变量又由另外的函数关系确定时,对于其测量结果的不确定度评定,本规范的基本原则也是适用的。但是评定起来比较复杂。

(3)对于被测量呈现为一系列值的分布,或对被测量的描述为一组量时,则测量结果的描述也应该是一组量值,测量不确定度应相应于每一个测量结果给出,并应给出与一组值相应的关系及分布情况。

(4)当被测量取决于一个或多个参变量时,例如以时间为参变量时,被测量的测量结果是随时间变化的直线或曲线,对于在直线或曲线上任意一点的测量结果和测量不确定度是不同的。测量不确定度的评定可能要用到最小二乘法、矩阵等数学运算,本规范在这方面没有详细规定,但本规范的基本原则也是适用的。

(5)本规范的基本原则也可用于对于统计控制下的测量过程的测量不确定度的评定,但评定时需要考虑测量过程的合并标准偏差作为 A 类标准不确定度。

第二章 基本术语 及其概念

测量不确定度的概念涉及到基本统计学术语及通用计量学术语。尤其是测量不确定度的采用导致了概率论、统计学和计量学方面的不少术语定义的修订：

(1) 在 1993 年 GUM 发布的同时，国际标准化组织发布了新版统计学术语：“ISO3534—1, 1993, Statistics—Vocabulary and Symbols—Part 1: Probability and General Statistics Terms” (ISO3534—1, 统计学 词汇和符号 第一部分：概率和通用统计学术语, 1993)。

(2) 在 1993 年 GUM 发布的同时，同样以 7 个国际组织的名义联合发布了 VIM 的修订版：“ISO(1993), International Vocabulary of Basic and General Terms in Metrology, Second edition” (国际通用计量学基本术语, 第二版, 1993)。

因此，在学习测量不确定度时，必须首先学习相关的概率统计方面的基础知识和计量学术语方面的基础知识，并了解在哪些方面与传统或经典的术语相比有了新的变化，由此加深我们对测量不确定度的理解。

我们可以发现，无论是概率论和数理统计的通用术语还是计量学通用术语都有两类：

一类是理论性的、概念性的术语，它们是在理想条件下的术语，在无穷多次测量条件下可以计算得到其极限值，但不可能实际测量得到。

另一类是实用性的术语，它是在实际使用中可以定量得到其量值的术语，即在有限次测量条件下可以通过测量或评定得到的，它们是概念性术语的估计值。

本章仅对与测量不确定度有密切关系的部分术语作简要的介绍。

第一节 基本的概率与数理统计术语及其概念

1. 概率(probability)

概率是一个 0 和 1 之间隶属于随机事件的实数。概率与在一段较长时间内的事件发生的相对频率有关，或与事件发生的可信程度(degree of belief)有关，在可信度高时概率接近 1。

在对某一个被测量重复测量时，我们可以得到一系列数据，这些数据称为测量值或观测值。测量值是随机变量，它们分散在某个区间内，概率是测量值在区间内出现的相对频率即可能性大小的度量；在此定义的基础上奠定了测量不确定度 A 类评定的理论基础。

由于测量的不完善或人们对被测量及其影响量的认识不足，会影响到测量值的可信度。

概率也可以是测量值落在某个区间内的可信度大小。在这新的定义中，对于那些我们不知道其大小的系统误差，可以认为是以一定的概率落在区间的某个位置，认为也属于随机变量，或者说某项未知的系统误差落在该区间内的可信程度也可以用概率表征。这是测量不确定度 B 类评定的理论基础。

以上两种情况都可认为是随机事件，这是对经典概率论的一个突破。

测量值 x 落在 $[a, b]$ 区间内的概率可以用 $P(a \leq x \leq b)$ 表示，概率也可简写为 P ，其值在 0 到 1 之间， $0 \leq P \leq 1$ 。

注：关于概率是一个事件发生相信程度的度量的观点参见 GUM 的 E.3 和 E.6。

2. 概率分布(probability distribution)

概率分布是一个随机变量取任何给定值或属于某一给定值集的概率随取值而变化的函数，该函数称概率密度函数。随机变量在整个值集的概率为 1。

(1)通俗地说，概率分布是单位区间内测量值出现的概率随测量值变化的分布情况。如图 2-1 所示，横坐标为测量值，纵坐标为概率密度函数 $p(x)$ 。

设 X 是连续随机变量， $p(x)$ 是 X 的概率密度函数。

若已知某个量的概率密度函数 $p(x)$ ，则测量值 X 的取值 x 落在区间 (a, b) 内的概率 P 可用下式计算：

$$P(a \leq X \leq b) = \int_a^b p(x) dx \quad (2-1)$$

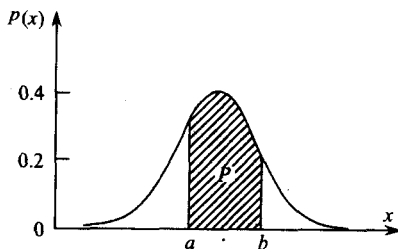


图 2-1 概率分布

(2)数学上，积分代表了面积。由此可见，概率 P 是概率分布曲线下在区间 (a, b) 内包含的面积。当 $P=0.9$ ，表明测量值有 90% 的可能性落在该区间内，该区间包含了概率分布下总面积的 90%。当 $P=1$ ，表明测量值的区间包含了概率分布下的全部面积，概率为 1 即测量值以 100% 的可能性落在该区间内，也就是测量值必定在此区间内。

为了与经典的“概率”有所区别，在 GUM 中将 P 称为“包含概率”或“置信水平(level of confidence)”，在 JJF1059 技术规范中也称为“置信概率”。

(3)在概率论中将具有一定置信概率的区间 $[a, b]$ 称为置信区间，在 GUM 中称“统计包含区间(statistical coverage interval)”。在概率论中通常用置信因子乘标准偏差($k\sigma$)得到置信区间的半宽度。在 GUM 中把为获得扩展不确定度(包含区间的半宽度 U)而用作合成标准不确定度的被乘因子称为包含因子($U=ku_c$)，也用符号 k 表示。置信区间的两个界限 a 和 b 分别称为区间的下限和上限，统称为置信限。通常，置信限位于测量结果最佳估计值的两侧，这样的区间称为双侧置信区间。但有时置信限仅位于一侧，即只有上限或只有下限，就称为单侧置信区间。大多数情况，置信限对称地位于测量结果最佳估计值的两侧，这样的区间称为对称置信区间。有些情况，置信限是不对称地位于测量结果最佳估计值的两侧的称不对称置信区间。

3. 期望(expectation)

期望又称概率分布或随机变量的均值(mean)或期望值(expected value)，有时又称数学期望。常用符号 μ 表示。

(1)对于离散随机变量测量值的期望是无穷多次测量的可能值 x_i 与其相应概率 p_i 乘积

之和, 设测量值 X 取值 x_i , 其概率为 P_i 时, 测量值的期望 $E(X)$ 为:

$$\mu = E(X) = \sum_{i=1}^{\infty} p_i x_i \quad (2-2)$$

(2) 对于连续随机变量若已知概率密度函数 $p(x)$ 时, 可按式计算测量值 X 的期望 $E(X)$:

$$\mu = E(X) = \int_{-\infty}^{+\infty} x p(x) dx \quad (2-3)$$

即通俗地说: 期望是无穷多次测量的平均值。

(3) 正因为实际上不可能进行无穷多次测量, 因此数学期望只可以通过数学运算得到, 但测量中是可望而不可得的。

(4) 期望是概率分布曲线与横坐标轴所构成面积的重心所在的横坐标, 所以期望是决定概率分布曲线位置的量。对于单峰、对称的概率分布来说, 期望值在分布曲线峰顶对应的横坐标。

(5) 进行无穷多次测量时, 测量值与其期望值之差为测量的随机误差, 测量值的期望值与真值之差就是测量的系统误差。真值是被测量的定义值。由此可见, 虽然真值、期望值和误差都是客观存在的, 但是, 都是理想条件下的概念, 因为不可能进行无穷多次测量, 并且真值未知, 也就不可能准确得到测量误差有多大。测量不可能没有误差, 因此不可能通过测量获得真值。

4. 方差 (variance)

随机变量或概率分布的方差是无穷多次测量随机误差平方的算术平均值的极限。用符号 σ^2 表示。

$$\sigma^2 = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (2-4)$$

(1) 测量值与期望值之差是随机误差, 用 δ 表示, $\delta_i = x_i - \mu$ 。方差就是随机误差的统计平均值。

(2) 测量值 X 的方差 $V(X)$ 是随机变量 X 的每一个可能值对其期望 $E(X)$ 的偏差的平方的期望。也就是测量的随机误差平方的期望。

$$\sigma^2 = V(X) = E\{[X - E(X)]^2\} \quad (2-5)$$

上式也可变换为如下形式:

$$\sigma^2 = V(X) = E(X^2) - E(X)^2 \quad (2-6)$$

已知测量值的概率密度函数时, 方差可表示为:

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 p(x) dx \quad (2-7)$$

当期望值为零时, 公式(2-8)可写成:

$$\sigma^2 = \int_{-\infty}^{+\infty} x^2 p(x) dx \quad (2-8)$$

(3) 方差说明了随机误差的大小和测量值分散程度, 但由于方差的量纲是单位的平方, 使用不方便、不直观, 因此引出了标准偏差这个术语。

5. 标准偏差 (standard deviation)

概率分布或随机变量的标准偏差是方差的正平方根值, 用符号 σ 表示。又称标准差。

$$\sigma = \sqrt{V(\bar{X})} \quad (2-9)$$

(1) 标准偏差是无穷多次测量的随机误差平方的算术平均值的正平方根值的极限。

$$\sigma = \lim_{n \rightarrow \infty} \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (2-10)$$

(2) μ 和 σ 对正态分布曲线的影响见图(2-2)。

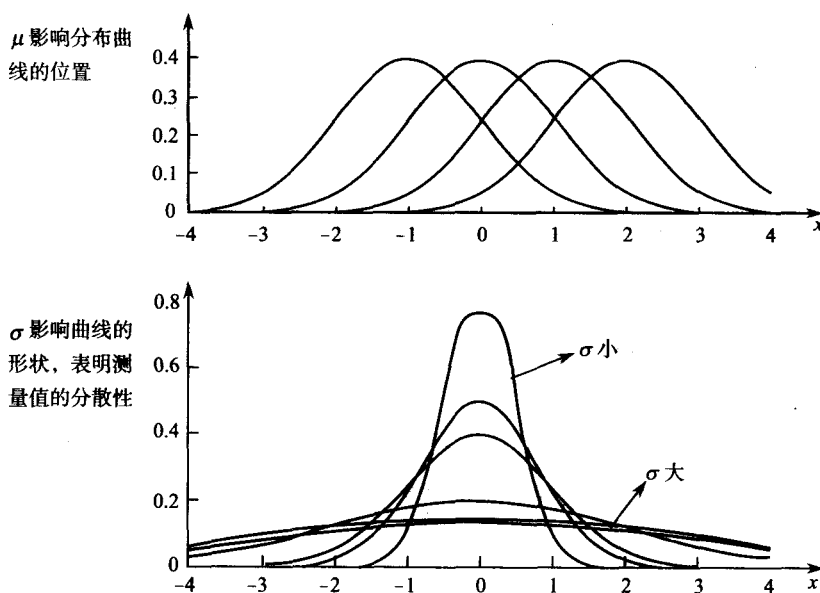


图 2-2 μ 和 σ 对正态分布曲线的影响

(3) 由图 2-2 可见, 标准偏差是表明测量值分散性的参数, σ 小表明测量值比较集中, σ 大表明测量值比较分散。通常测量的重复性或复现性是用标准偏差 σ 的估计值来表示的。

(4) 由于标准偏差 σ 是无穷多次测量时的极限值, 所以又称总体标准差。

可见: 期望和方差(或标准偏差)是表征概率分布的两个特征参数。理想情况下, 应该以期望为测量结果, 标准偏差表示测量值的分散性。

(5) 由于期望、方差和标准偏差都是以无穷多次测量的理想情况定义的, 因此都是概念性的术语, 无法由测量得到 μ , σ^2 和 σ 。

6. 算术平均值(arithmetic mean) —— 期望的最佳估计值

在相同条件下对被测量 X 进行有限次独立重复测量, 得到一系列测量值 $x_1, x_2, \dots, x_n$, 其算术平均值 $\bar{X}$ 为:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2-11)$$

① 由大数定理证明, 测量值的算术平均值是其期望的最佳估计值。

大数定理: 设 ϵ 为无限小量, 则 $\lim_{n \rightarrow \infty} P\left\{\left|\left(\frac{1}{n} \sum_{i=1}^n x_i\right) - \mu\right| < \epsilon\right\} = 1$ 即 $\lim_{n \rightarrow \infty} P(|\bar{X} - \mu| < \epsilon) = 1$ 。

说明若干个独立同分布的随机变量的平均值以无限接近于 1 的概率接近于其期望值 μ 。所以

$\bar{X}$ 是期望 μ 的最佳估计值。

②算术平均值是有限次测量的均值，所以是由样本构成的统计量。即使在同一条件下对同一量进行多组测量，每组的平均值都不相同，说明算术平均值本身也是随机变量。

③由于测量值的算术平均值是其期望的最佳估计值，因此，通常用算术平均值作为测量结果。

7. 实验标准偏差 (experimental standard deviation) ——有限次测量时标准偏差的估计值

实际工作中不可能测量无穷多次，因此无法得到总体标准偏差 σ 。用有限次测量的数据得到的标准偏差的估计值称为实验标准偏差，用符号 s 表示。以下介绍几种常用的实验标准偏差的估计方法。

在相同条件下，对被测量 X 作 n 次独立重复测量，每次测量值为 x_i ，测量次数为 n ，则实验标准偏差可按以下几种方法估计：

(1) 贝塞尔公式法

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}} \tag{2-12}$$

式中： $\bar{X}$ —— n 次测量的算术平均值；

$v_i = x_i - \bar{X}$ ——残差；

$\nu = n - 1$ ——自由度；

s ——实验标准偏差。

残差是测量值与算术平均值之差。测量值的误差不可能通过测量得到，但残差是可以获得的。由贝塞尔公式估计的标准偏差是被测量残差的统计平均值。

自由度 (degrees of freedom) 是指计算残差平方和时具有独立项的个数。因为 n 较大时，残差和为零，因此 n 个残差中任何一个残差可以从另外 $n-1$ 个残差中推算出来，独立的残差项只有 $n-1$ 个，也就是自由度为 $n-1$ 。可理解为：被测量只有一个时，为估计被测量，只需测量一次，但为了提高测量的可信度而多测了 $n-1$ 次，多测的次数可以酌情规定，所以称自由度。由此可以推论，当待测量为 t 个，测量次数为 n 时，则自由度为 $n-t$ ；如果另有 r 个约束条件，则自由度为 $n-t-r$ 。所以自由度通常情况下定义为“总和的项数减去总和中受约束的项数”。在给出标准偏差的估计值时，最好同时给出其自由度，自由度越大，表明估计值的可信度越高。

(2) 最大残差法

从有限次独立重复测量的一系列测量值中找出最大残差 $v_{\max}$ ，并根据测量次数 n 查表 2-1 得到 c_n 值，代入下式得到估计的标准偏差：

$$s = c_n |v_{\max}| \tag{2-13}$$

表 2-1 最大残差法的 c_n 值表

n	2	3	4	5	6	7	8	9	10	15	20
c_n	1.77	1.02	0.83	0.74	0.68	0.64	0.61	0.59	0.57	0.51	0.45