

山东省建设工程质量检测能力验证技术导则

Technical Guidelines on Proficiency Testing for Construction Engineering

Quality Inspection in Shandong Province

山东省住房和城乡建设厅

2022年6月

前言

为进一步提高我省建设工程质量检测水平，规范建设工程质量检测能力验证工作，山东省住房和城乡建设厅组织有关单位经调查研究，认真总结实践经验，参考有关国家标准，并在广泛征求意见的基础上，制定本导则。

本导则的主要技术内容：1.总则；2.术语；3.基本规定；4.能力验证的组织者；5.能力验证的承担者；6.能力验证的参加者；7.能力验证样品；8.能力验证结果统计与能力评价；9.能力验证结果处理；10.能力验证服务管理平台。

本导则由山东省住房和城乡建设厅负责管理，由山东省建设工程质量安全中心负责具体条文内容的解释。执行过程中如有意见和建议，请寄送山东省建设工程质量安全中心（地址：济南市历下区朝山街 25 号，邮政编码：250011，电话：0531-51765311，电子邮箱：sdzljdz_zy@shandong.cn）。

主 编 单 位：山东省建设工程质量安全中心

参 编 单 位：聊城市建设工程质量服务保障中心

山东省建筑科学研究院有限公司

济南市工程质量与安全中心

荣成市建筑工程事务服务中心

主要起草人员：张 毅 王桂珍 吴 军 边 琦 王 志

卢晓芳 许 震 戚爱谦

主要审查人员：徐新生 张 霞 夏攀登 刘巧玲 张省祥

刘 聪 向守元

目录

1 总则.....	1
2 术语	2
3 基本规定	4
4 能力验证的组织者	5
5 能力验证的承担者	7
5.1 一般规定.....	7
5.2 人员.....	8
5.3 设备、设施和环境.....	8
5.4 能力验证实施方案.....	9
5.5 能力验证的运作.....	10
5.6 技术报告.....	10
6 能力验证的参加者	12
7 能力验证样品	13
8 能力验证结果统计与能力评价	15
8.1 一般规定.....	15
8.2 统计方法的选择.....	15
8.3 指定值及其不确定度的确定.....	16
8.4 能力验证结果的统计处理.....	16
8.5 能力统计量的计算.....	17
8.6 能力评定.....	18
9 能力验证结果处理	20
10 能力验证服务管理平台	21
附录 A 能力验证实施详细流程图.....	22
附录 B 能力验证实施方案.....	23
附录 C 能力验证作业指导书.....	31
附录 D 能力验证技术报告.....	32
附录 E 样品的均匀性和稳定性检验	37
附录 F 检测能力验证样品确认单	43
附录 G 结果统计分析方法.....	44
本导则用词说明	57
引用标准名录	58
附：条文说明	59

Contents

1	General Provision.....	1
2	Terms	2
3	Basic Requirements	4
4	Proficiency Testing Organizer	5
5	Proficiency Testing Commitment Organization	7
5.1	General Requirement	7
5.2	Personnel.....	8
5.3	Equipment, Facilities and Environment	8
5.4	Implementation Plan of Proficiency Testing.....	9
5.5	Operation of Proficiency Testing	10
5.6	Technical Report	10
6	Proficiency Testing Participant.....	12
7	Proficiency Testing Item.....	13
8	Statistics of Proficiency Testing Results and Performance Evaluation	15
8.1	General Requirement	15
8.2	Selection of statistical methods.....	15
8.3	Determination of Assigned Value and Uncertainty	16
8.4	Statistical Treatment of Proficiency Testing Results	16
8.5	Calculation of Proficiency Testing statistics	17
8.6	Performance Evaluation	18
9	Treatment of Proficiency Testing Results.....	20
10	Proficiency Testing Service and Management Platform	21
Appendix A	Detailed Flow Chart of Proficiency Testing Implementation	22
Appendix B	Implementation Plan of Proficiency Testing	23
Appendix C	Work Instruction of Proficiency Testing.....	31
Appendix D	Technical Report on Proficiency Testing.....	32
Appendix E	Test on the Homogeneity and Stability of Samples.....	37
Appendix F	Proficiency Testing Sample Confirmation Form	43
Appendix G	Statistical Analysis Method of the Results.....	44
	Explanation of Wording in the Guidelines	57
	List of Quoted Standards	58
	Addition:Explanation of Provisions.....	59

1 总 则

1.0.1 为加强山东省建设工程质量检测管理，规范建设工程质量检测能力验证工作，确保能力验证工作的科学性、规范性、有效性，保证检测机构技术能力持续符合要求，制定本导则。

1.0.2 本导则适用于在山东省行政区域内建设工程质量检测能力验证工作的组织实施过程控制和结果利用的管理。

1.0.3 建设工程质量检测能力验证除应符合本导则外，尚应符合国家、行业及地方相关标准、规范和管理规定。

2 术 语

2.0.1 建设工程质量检测 testing for quality of construction engineering

按照相关规定的要求，采用试验、测试等技术手段确定建设工程的建筑材料、工程实体质量特性的活动。

2.0.2 实验室间比对 inter-laboratory comparison

按照预先规定的条件，由两个或多个实验室对相同或类似的样品进行测量或检测的组织、实施和评价。

2.0.3 能力验证 proficiency testing

利用实验室间比对，按照预先确定的判据评价参加者的能力。

2.0.4 能力验证的组织者 proficiency testing organizer

负责能力验证统一协调、组织实施、监督管理的机构。

2.0.5 能力验证的承担者 proficiency testing commitment organization

受组织者委托，承担能力验证具体实施工作的机构。

2.0.6 能力验证的参加者 proficiency testing participant

参加能力验证的建设工程质量检测机构。

2.0.7 能力验证样品 proficiency testing sample

用于能力验证的样品、产品、人工制品、标准样品、设备部件等。

2.0.8 稳健统计方法 robust statistical method

对给定的概率模型和假定条件下，对少量偏离数据不敏感的统计方法。

2.0.9 总计统计量 total statistics

结果总数、中位值、稳健变异系数、极大值、极小值、变动范围、标准化四分位距七个统计量在稳健统计技术中称之为总计统计量。

2.0.10 指定值 assigned value

对能力验证样品的某个特性赋予的值。

2.0.11 公议值 consensus value

由实验室间比对的一组结果得到的值。

2.0.12 离群值 outlier

一组数据中被认为与该组其他数据不一致的观测值。

2.0.13 能力评定标准差 standard deviation for proficiency assessment

根据可获得的信息，用于评价能力验证结果分散性的度量。

2.0.14 不确定度 uncertainty

表征合理地赋予被测量之值的分散性，与测量结果相联系的参数。

2.0.15 z 比分数 z-score

使用指定值和能力评价标准差计算的实验室偏倚的标准化能力统计量。

注：z 比分数也称为 z 值或 z 分数，z 也可为大写。

2.0.16 满意结果 satisfactory results

通过能力验证活动，利用统计技术或专家公议等技术手段，确认参加者的能力为满意的结果。

2.0.17 可疑结果 questionable results

通过能力验证活动，利用统计技术或专家公议等技术手段，确认参加者的能力可能出现问题的结果。

2.0.18 不满意结果 unsatisfactory results

通过能力验证活动,利用统计技术或专家公议等技术手段，确认参加者的能力为不满意的结果。

2.0.19 能力验证服务管理平台 proficiency testing service and management platform

利用计算机技术、网络通信技术等信息化手段，实现能力验证信息进行发布、样品确认、数据接收、结果评定等功能的网络管理系统。

3 基本规定

- 3.0.1** 检测能力验证工作应当遵循客观公正、科学合理、统一规范的原则。
- 3.0.2** 能力验证的组织者可委托有能力的机构承担本行政区域内的检测能力验证工作。
- 3.0.3** 能力验证的承担者应当按照能力验证的组织者制定的能力验证基本规范和实施规则开展能力验证活动。
- 3.0.4** 能力验证的承担者应当满足开展能力验证工作的要求，其技术能力在相应领域和关键技术要素方面领先，并具备可持续性。能力验证的组织者负责对能力验证的承担者进行确认。
- 3.0.5** 能力验证的参加者应按规定参加检测能力验证活动。
- 3.0.6** 能力验证的组织者负责能力验证的组织实施，能力验证的承担者负责相应实施方案的设计、样品的提供及检测结果统计分析，能力验证的参加者负责按计划要求完成检测任务及实施相关改进措施。能力验证实施流程见附录 A。
- 3.0.7** 能力验证的组织者可组织能力验证结果可疑或不满意的参加者进行二次能力验证，程序与初次能力验证相同。

4 能力验证的组织者

4.0.1 能力验证的组织者包括省、设区市、县三级住房城乡建设主管部门。省住房城乡建设主管部门负责统一协调、组织实施、监督管理全省检测能力验证工作。设区市、县级住房城乡建设主管部门负责本行政区域内能力验证的统一协调、组织实施和监督管理工作。

4.0.2 能力验证的组织者负责以下工作（包括但不限于）：

- 1 对能力验证进行全面管理，下达年度能力验证计划通知；
- 2 确定能力验证的参加者名单；
- 3 对能力验证的承担者进行考查、甄选；
- 4 审核能力验证实施方案、能力验证技术报告；
- 5 发布能力验证结果通报及相关结果处置规定；
- 6 组织能力验证结果分析及技术交流会议；
- 7 组织可疑或不满意结果的检测机构进行整改并验证整改效果；
- 8 能力验证项目验收。

4.0.3 能力验证的组织者根据检测管理工作需要，制定年度能力验证工作计划。制定计划时，应当优先考虑涉及结构安全、节能环保、主要使用功能等领域的的能力验证项目。

4.0.4 能力验证的组织者应根据能力验证项目和取得相应资质证书的检测机构名单，确定能力验证的参加者名单，通知其参加相应的能力验证活动。

4.0.5 能力验证的组织者应对能力验证的承担者进行考查、甄选，确保其能力满足能力验证项目的组织和技术要求。

4.0.6 能力验证的组织者应与能力验证的承担者明确双方权利和义务，并对承担者及其承担的能力验证活动进行监督，发现承担者在能力验证工作中存在重大问题或者能力验证结果评价不合理等情形的，应当及时督促其改正。

4.0.7 能力验证的组织者应对承担者提供的能力验证方案、样品均匀性和稳定性检验结果、能力验证技术报告进行审核，及时纠正存在的问题，并监督承担者对相应内容作进一步修改和完善。

4.0.8 能力验证的组织者应及时向社会公布其组织实施的能力验证结果信息。

4.0.9 能力验证的组织者对于能力验证出现可疑或不满意结果的检测机构，应要求其在规定期限内进行整改并验证整改效果。

4.0.10 对能力验证结果存在异议的，能力验证的参加者可以在收到能力验证结果之日起 15 个工作日内向能力验证的组织者提出申诉。能力验证的组织者应组织能力验证的承担者对申诉内容进行研究，并及时给出答复。

4.0.11 能力验证的组织者应建立并保存能力验证档案及相关记录，包括：

- 1 实施能力验证的有关文件；

- 2 能力验证承担者的相关资质证明及对其确认的记录；
- 3 能力验证参加者名单；
- 4 能力验证的技术报告；
- 5 能力验证结果和后续处理相关文件。

5 能力验证的承担者

5.1 一般规定

5.1.1 能力验证的承担者应按照相关技术标准的规定实施能力验证活动：策划、制定能力验证方案；制备能力验证样品，并对其进行验证、定值、分发；对参加能力验证的检测机构提交的数据进行统计分析和结果评价，编制能力验证技术报告。

5.1.2 能力验证的承担者应符合以下要求：

- 1 依法成立并能够承担相应法律责任的法人或经法人授权的其他组织；
- 2 具有相关领域的技术专家和统计专家等人员，能够按照相应标准或技术要求制定能力验证方案、开展结果统计评定，保障能力验证工作的有效性；
- 3 具有承担能力验证活动所必需的设备、设施和环境；
- 4 具有并有效运行保证其能力验证活动规范、独立、公正、科学、诚信的管理体系；
- 5 能够制备或者取得能力验证样品，保证能力验证样品均匀性和稳定性，并对能力验证样品进行有效管理，包括样品存储、包装、标识、分发和处置；
- 6 能够客观、清晰、全面地编制能力验证技术报告，技术报告应包含所有参加机构的能力验证结果等内容；
- 7 未被列入经营异常名录或严重违法失信名单。

5.1.3 能力验证的承担者按照组织者的要求编写能力验证实施方案，依据组织者的意见完善方案后，开始能力验证样品的准备工作。能力验证实施方案格式见附录 B。

5.1.4 能力验证的承担者应编写能力验证作业指导书并随样品发送能力验证的参加者，指导参加者正确接收和处置样品、正确实施检测和填报检测结果等。能力验证作业指导书格式见附录 C。

5.1.5 能力验证的承担者承担的能力验证样品的准备和发放、统计分析和结果评价等工作，应符合本导则的要求。

5.1.6 能力验证的承担者在能力验证工作完成后，应及时向能力验证的组织者报送能力验证技术报告，内容包括：能力验证项目名称、项目实施起止时间、验证的检测参数、依据的检测标准、能力验证样品的均匀性和稳定性检验结果、参加机构名单和评价结果、统计数据、技术分析和建议等信息，由组织者对项目进行验收。

5.1.7 能力验证的承担者应准确、客观、公正地实施能力验证活动，并对所出具的统计结果、评价结论、能力验证结果和报告负责。

5.1.8 能力验证的承担者应保存能力验证活动的原始记录、数据信息和结果报告，保存期限不少于 6 年。

5.2 人员

5.2.1 能力验证的承担者中的人员应具备以下能力：

- 1 选择适当的能力验证样品；
- 2 策划能力验证方案；
- 3 进行特定类型的抽样；
- 4 操作特定的仪器设备；
- 5 进行测量以确定能力验证样品被测量的稳定性和均匀性、指定值及其不确定度；
- 6 制备、处置并分发能力验证样品；
- 7 操作数据处理系统；
- 8 进行统计分析；
- 9 评价能力验证参加者的能力；
- 10 提供技术意见和解释；
- 11 批准发布能力验证技术报告。

5.2.2 能力验证承担者应为实施能力验证工作的技术人员制定教育和培训计划，确保人员获得必要的培训，保证人员能够胜任检测、仪器设备操作以及其他影响能力验证质量的工作，对人员进行评价和授权，并保留所有人员的相关授权、能力、教育和培训等的最新记录。

5.3 设备、设施和环境

5.3.1 能力验证承担者应具有与实施能力验证工作相适应的设施，主要包括用于能力验证样品的制备、处置、检测、储存和分发，以及数据处理、通讯、资料和记录检索的设施和设备。

5.3.2 能力验证承担者应确保环境条件不会影响能力验证工作的实施质量。当能力验证工作是在能力验证承担者固定设施以外的场所或委托给分包方实施时，应将影响能力验证工作的设施和环境条件的技术要求制定成文件。

5.3.3 能力验证承担者应对影响能力验证工作质量的区域的进入和使用加以控制。

5.3.4 能力验证承担者应识别可能显著影响能力验证样品或检测质量的环境条件，应对这些条件加以控制和监测，并记录所有相关监控活动。当环境条件危及到能力验证工作的质量或实施时，应停止相关能力验证活动。

5.3.5 应将不相容活动的相邻区域进行有效隔离，并采取措施以防止交叉污染。

5.4 能力验证实施方案

5.4.1 能力验证承担者应识别和策划直接影响能力验证工作质量的过程，并确保这些过程依据既定程序进行。

5.4.2 能力验证承担者不得将能力验证的策划工作分包，但可向顾问、专家或指导小组寻求建议和帮助。

5.4.3 能力验证实施方案应包括以下内容：

- 1 本次能力验证的目标、目的以及基本设计情况；
- 2 能力验证承担者的名称和地址；
- 3 协调者或其他参与能力验证设计和实施人员的姓名、地址和联系方式；
- 4 分包的活动以及参与能力验证实施分包方的名称和地址；
- 5 参加能力验证应满足的条件；
- 6 能力验证的参加者数量和类型；
- 7 所选定的被测量或特性，包括参加者需要鉴别、测量或检测的有关信息；
- 8 对能力验证样品预期的量值范围和/或特性的描述；
- 9 所提供能力验证领域中涉及的潜在的主要错误来源；
- 10 对能力验证样品生产、质量控制、存储、分发的要求；
- 11 合理防范参加者串通或伪造结果的措施，以及当怀疑串通或伪造时可执行的程序；
- 12 将提供给参加者的信息描述，以及能力验证各阶段时间表；
- 13 当连续能力验证时，给参加者分发能力验证样品的频次或日期，参加者返回结果的截止日期，适当时，参加者进行检测或测量的日期；
- 14 参加者准备检测材料以及进行检测所使用的方法或程序的有关信息；
- 15 用于能力验证样品均匀性和稳定性检验的检测方法的程序；
- 16 为参加者准备的标准化的结果报告格式；
- 17 所使用统计分析的详细描述；
- 18 所有指定值的来源、计量溯源性和测量不确定度；
- 19 参加者能力评价的准则；
- 20 返回给参加者的数据、中期报告或信息的描述；
- 21 参加者结果和根据能力验证结果所做结论的公布范围描述；
- 22 能力验证样品丢失或损坏时应采取的措施。

5.4.4 能力验证承担者可依据能力验证实施方案的设计，要求参加者使用指定的方法。若允许参加者使用他们自己选择的方法，能力验证提供者应制定政策并按程序对不同检测方法得到的结果进行比对，了解不同检测方法的技术等效性，并采取措施评价使用不同方法的参加者的结果。

5.5 能力验证的运作

5.5.1 能力验证承担者应在发放能力验证样品前及早告知参加者能力验证样品可能到达或将要分发的日期，除非能力验证实施方案的设计不适宜这样做。

5.5.2 能力验证承担者应向所有参加者提供详细的文件化的作业指导书。指导书应包括：

- 1 要求参加者按照日常检测样品的处理方式处理能力验证样品；
- 2 对于影响能力验证样品检测因素的详细说明，例如：能力验证样品的性质、存储条件、是否限定检测方法，以及检测的时间要求等；
- 3 进行检测之前，能力验证样品的准备和/或状态调节的详细程序；
- 4 处置能力验证样品的适当指导，包括安全要求；
- 5 参加者检测时特定的环境条件，如适用，要求参加者报告检测期间相关环境条件；
- 6 检测结果及其不确定度记录和报告方式的明确和详细的说明。如果指导书要求报告结果的测量不确定度，应包括包含因子和置信概率（只要可行）；
- 7 能力验证承担者接收用于分析的能力验证结果的截止日期；
- 8 能力验证承担者接受问询的详细联络信息；
- 9 适用时，返回能力验证样品的说明。

5.6 技术报告

5.6.1 能力验证技术报告应清晰、全面，包含所有参加者结果的资料，并指出每个参加者的能力。最终报告的批准不应分包。

5.6.2 能力验证技术报告应包括以下内容：

- 1 能力验证相关方的名称和详细联系信息；
- 2 报告批准人的姓名、职位、签名或等效标识；
- 3 承担者分包活动的说明；
- 4 报告发布日期和状态（如初期的、中期的或最终的）；
- 5 报告的页码和清晰的结束标记；
- 6 能力验证技术报告的编号和清晰标识；
- 7 对能力验证样品的清晰描述，包括能力验证样品制备、均匀性和稳定性评定的必要细节；
- 8 参加者的结果；
- 9 统计数据及总计统计量，包括指定值、可接受结果的范围和图形表示；
- 10 用于确定指定值的程序；
- 11 指定值的测量不确定度的详细信息；
- 12 用于确定能力评定标准差或其他评定准则的程序；

13 对应每组参加者使用的检测方法/程序的指定值和总计统计量（如果不同组的参加者使用了不同的方法）；

14 能力验证承担者和技术顾问（如有）对参加者的能力评述；

15 能力验证实施方案的设计和实施的信息；

16 数据统计分析的程序；

17 对统计分析解释的建议；

18 能力验证结果的评述或建议。

能力验证技术报告格式见附录 D。

6 能力验证的参加者

- 6.0.1** 能力验证的参加者应按照能力验证组织者的要求参加相应能力验证活动。
- 6.0.2** 能力验证的参加者应对参加能力验证的过程进行严格的控制和管理。在能力验证试验开始前，参加者应组织技术人员核查所用仪器设备、标准物质、环境条件是否符合能力验证的要求，认真研究检测方法，制定检测方案。参加者可组织检测人员模拟能力验证试验过程，自行制备样品，进行内部比对，提前解决模拟过程中发现的问题，为能力验证试验做好充分准备。
- 6.0.3** 能力验证的参加者应当依据作业指导书和相关标准的要求，独立完成能力验证样品检测，并在规定时间内真实、客观地报送检测数据、结果及相关原始记录，不得私下比对串通能力验证数据、结果或出具虚假能力验证数据、结果。
- 6.0.4** 能力验证的参加者应当向组织者及时反馈相关信息，并保存相关记录。
- 6.0.5** 能力验证结果可疑或不满意的，能力验证的参加者应进行原因调查，采取相应的纠正措施，并根据组织者的要求参加二次能力验证。
- 6.0.6** 能力验证的参加者应当保存能力验证活动的原始记录、数据信息和结果报告，保存期限不少于 6 年。

7 能力验证样品

7.0.1 能力验证的承担者应正确制备与所开展能力验证项目相适应的样品，确定样品均匀性、稳定性，并采取必要的控制措施，对样品特性值和相关信息进行保密，确保样品从制备到分发的过程中得到妥善标识、隔离和免受污染、混淆或受损。

7.0.2 能力验证样品的选择和制备应遵循以下原则：

1 能力验证项目选择的被测样品，在性质上应与参加者的日常检测样品或材料相类似，在选用标准物质作为样品时，应谨防被能力验证的参加者猜出结果；

2 应选择预期均匀性、稳定性较好的材料制备能力验证样品；

3 能力验证的样品制备可由能力验证的承担者自己进行，也可外包，当选择外包方时应考察外包方是否具有制备该样品的能力，包括设备（满足生产小批量的能力验证样品的要求）、生产工艺的技术要求、人员等方面是否满足要求；

4 被测样品的数量应满足参加能力验证的需要，包括分发给参加者的样品（包括二次能力验证）、用于均匀性、稳定性检验的样品，还应考虑制备额外数量的被测样品，以替代样品分发过程中丢失或损坏的样品。

7.0.3 样品均匀性、稳定性试验和评价应符合以下要求：

1 样品为承担者制备的，制备完成后，能力验证的承担者按照本导则附录 E 的要求进行样品均匀性、稳定性试验和评价。

2 样品为购置的，应从样品提供单位取得均匀性、稳定性的证明，并对均匀性和稳定性进行验证。

3 不需要进行稳定性试验的应有充分的依据并进行说明。

4 当无法进行均匀性和稳定性检验时，能力验证的承担者应证明能力验证样品的收集、制备、包装和分发程序可以充分满足能力验证要求。

7.0.4 能力验证的承担者在完成样品均匀性、稳定性确认后，应报能力验证的组织者进行审核，确认样品满足要求后，由承担者包装和发送样品。

7.0.5 能力验证的承担者应当采用适当的包装和运输方式进行样品发放，保证被检测样品的稳定性和特性。应确保标识牢固地贴在每个能力验证样品的包装上，并在整个能力验证期间保持清晰和完整。

7.0.6 能力验证的承担者分发样品的同时，应要求参加者填写《检测能力验证样品确认单》（见附录 F）。如发现样品有损坏或其他异常情况，能力验证的承担者应及时采取补救措施。

7.0.7 能力验证的参加者应严格按照作业指导书的要求，对能力验证样品进行保管、制备、检测，不得将能力验证的样品分包给另一个检测机构进行检测。

7.0.8 需要时，能力验证样品在试验后返回能力验证的承担者，能力验证的参加者宜采用与发送样品同等要求的包装。

8 能力验证结果统计与能力评价

8.1 一般规定

- 8.1.1** 能力验证承担者应运用适当的方法记录和分析参加者提交的结果。
- 8.1.2** 能力验证承担者应建立和执行程序以检查数据输入、转换、统计分析和报告的有效性。
- 8.1.3** 数据分析应给出与能力验证统计设计相符的总计统计量、能力统计量以及有关信息。
- 8.1.4** 对能力验证结果统计分析能力及能力评价一般包括以下几方面内容：
- 1 统计方法的选择；
 - 2 指定值的确定；
 - 3 能力验证结果的统计处理；
 - 4 能力统计量的计算；
 - 5 能力评定。
- 结果统计分析方法见附录 G。
- 8.1.5** 必要时，需考虑能力验证样品的均匀性和稳定性对能力评定的影响。

8.2 统计方法的选择

- 8.2.1** 能力验证承担者应根据数据的特性、统计假定、误差的性质以及预期的结果数量，选择符合能力验证目标的统计方法。
- 8.2.2** 能力验证承担者应对选用统计方法的原因和假定进行说明，并能证实统计假定合理。
- 8.2.3** 选择统计方法时，应考虑以下事项：
- 1 能力验证每个被测量或特性所要求或期望的准确度（正确度和精密度）以及测量不确定度；
 - 2 达到统计设计目标所需的最少参加者数量。当参加者数量不足以达到目标或不能对结果进行有效的统计分析时，能力验证承担者应制定能力评定的替代方法，并形成文件，提供给参加者；
 - 3 有效数字和/或小数位数与所报告结果的相关性；
 - 4 需要检测的能力验证样品数量，以及对每个能力验证样品或每项测定的检测重复次数；
 - 5 用于确定能力评定标准差或其他评定准则的程序；
 - 6 用于识别和（或）处理离群值的程序；

- 7 对统计分析中剔除值的评价程序；
- 8 与设计相符的目标和能力验证轮次的频率。

8.3 指定值及其不确定度的确定

8.3.1 指定值的确定有多种方法，以下列出最常用的方法：

- 1 已知值——根据特定能力验证样品配方（如制造或稀释）确定的结果。
- 2 有证参考值——根据定义的检测或测量方法确定（针对定量检测）。若能力验证样品为有证标准物质（CRM），则该物质的参考值就可作为指定值。当能力验证样品为有证标准物质时，其指定值的标准不确定度由证书中给出。
- 3 参考值——根据对能力验证样品和可溯源到国家标准或国际标准的标准物质/标准样品或参考标准的并行分析、测量或比对来确定。
- 4 专家公议值——以一次比对试验前或试验中一组专家实验室报告结果的稳健平均值得到。
- 5 参加者确定的公议值——由某轮能力验证参加者（预先确定的比例，如 80%或更高）报告的结果计算得出。

8.3.2 如果指定值由参加者公议确定，应当有确定该指定值正确度和检查数据分布的程序。例如，可采用将指定值与一个具备专业能力的实验室得到的参考值进行比较等方法确定指定值的正确度。

8.3.3 如果符合以下准则，则可认为指定值的不确定度可以忽略不计，不需要在该轮能力验证结果中予以解释。指定值不确定度的确定参照 GB/T 28043-2019 第 7.2 条。

$$u(x_{pt}) < 0.3\sigma_{pt}$$

式中： $u(x_{pt})$ 为稳健平均值的标准不确定度；

σ_{pt} 为稳健标准差。

注：1. $u(x_{pt})$ 计算见附录公式 G.7。

2. 基于实用性考虑，本导则未给出不符合以上准则时需要采取的措施，如有必要，可参考 GB/T28043-2019 第 9 章。

8.4 能力验证结果的统计处理

8.4.1 能力验证的承担者应根据能力验证所要求遵从的检测方法标准和统计分析的要求，明确规定参加者在上报每个检测结果及计算平均值时所应该获取的有效位数及其所应执行

的数字修约标准（如 GB/T 8170 等）。

8.4.2 能力验证的承担者应在预定日期内收集参加者的结果报告单。对于逾期上报或存在明显错误的的数据，可不计入统计中，但应在能力验证技术报告中予以说明。

8.4.3 若结果分布为近似正态分布，则可用稳健技术分析；若结果分布为不对称分布，则先用核密度图识别多峰性，并使用适用于不对称结果的稳健技术。

8.4.4 在数据分析的早期阶段，即使用稳健方法或通过检验来识别统计离群值之前，应从数据集内删除明显错误的的数据。通常，对这些结果将分别予以处理（比如联系相关参加者）。这种方法可以纠正某些错误，但应依照批准的政策和程序进行处理。

能力验证承担者应安排具备充分专业技术和统计知识的人员对数据进行直观检查，以确认结果的预期分布，识别异常值或意外变异的来源。

8.4.5 离群值可按下列方法进行统计处理：

1 明显错误的结果应由专家进行识别和判断。如单位错误、小数点错误、计算错误或者错报为其他能力验证样品的结果，应从数据集中剔除，单独处理。这些结果不再计入离群值检验或稳健统计分析。

2 当使用参加者的结果确定指定值时，应使用适当的统计方法使离群值的影响降到最低，即可以使用稳健统计方法或计算前剔除离群值。

3 如果某结果作为离群值被剔除，则仅在计算总计统计量时剔除该值。但这些结果仍应当在能力验证中予以评价，并进行适当能力评定。

8.4.6 能力验证的设计应考虑到满足设计目标所需要的参加者的最小数量，并说明如果不能达到该最小数量时应使用的替代方法。少量参加者的能力验证统计方法可参考附录 G 第 G.3 条。

8.5 能力统计量的计算

8.5.1 定量结果的常用统计量如下：

1 偏差（测量误差）或百分相对差

$$D_i = x_i - x_{pt} \quad (1)$$

$$D_i\% = \frac{(x_i - x_{pt})}{x_{pt}} \times 100 \quad (2)$$

式中 x_i 为参加者*i*报告的结果， x_{pt} 为指定值。如果用 δ_E 表示测量结果的最大允许误差，当 $-\delta_E < D < \delta_E$ ，则认为能力评定是“可接受”的（根据 δ_E 的表达式，相同的准则也可应用于 $D\%$ ）。

2 z 值

$$z_i = \frac{x_i - x_{pt}}{\sigma_{pt}} \quad (3)$$

式中 x_i 为参加者*i*报告的结果， x_{pt} 为指定值， σ_{pt} 为能力评定标准差。

确定能力评定标准差通常有五种方法：规定值、经验值、一般模型、测量方法精密度和由参加者结果确定。

3 z' 值

$$z'_i = \frac{x_i - x_{pt}}{\sqrt{\sigma_{pt}^2 + u^2(x_{pt})}} \quad (4)$$

当指定值不确定度 $u(x_{pt}) \geq 0.3\sigma_{pt}$ 时，能力评定标准差中需考虑加入指定值不确定度分量。 z 值的评定标准适用于 z' 值。

4 ζ 值

$$\zeta_j = \frac{x_i - x_{pt}}{\sqrt{u^2(x_i) + u^2(x_{pt})}} \quad (5)$$

式中 $u(x_i)$ 是参加者结果 x_i 的标准不确定度， $u(x_{pt})$ 是指定值 x_{pt} 的标准不确定度。

5 E_n 值

$$(E_n)_i = \frac{x_i - x_{pt}}{\sqrt{U^2(x_i) + U^2(x_{pt})}} \quad (6)$$

式中 x_{pt} 为参考实验室确定的指定值， $U(x_i)$ 为参加者结果 x_i 的扩展不确定度， $U(x_{pt})$ 为 x_{pt} 的扩展不确定度。

8.6 能力评定

8.6.1 应根据能力度量方式制定能力评定准则，用于能力评定的方式如下：

- 1 与目标的符合性，根据方法性能指标和参加者的操作水平等预先确定准则。
- 2 用统计方法确定的 z 比分数和 z' 比分数，具体准则如下：

z 值的评定标准为：

当 $|z| \leq 2.0$ 时，表示满意结果；

$2.0 \leq |z| \leq 3.0$ 时，表示可疑结果；

$|z| \geq 3.0$ 时，表示不满意结果。

注：上述 z 比分数的评定准则，也适用于 z' 比分数。

E_n 值的评定标准为：

当 $|E_n| \leq 1.0$ 时，表示结果满意；

$|E_n| > 1.0$ 时，表示结果不满意。

3 采用专家公议的方式，由技术专家直接确定报告结果是否与预期目标相符合。如混凝土结构实体强度检验能力验证可采用此方式。

8.6.2 能力验证的承担者可利用 z 比分数（ z ， z_b 和 z_w ）序列柱状图、直方图和 Youden

（尧敦）图等图形来比对和分析各参加者的检测能力。

8.6.3 通过参加者结果与指定值之差完全可以确定参加者的能力，差值 D 也称为“实验室偏倚的估计值”；百分相对差不依赖于指定值的大小，参加者也很容易理解。

8.6.4 ζ 值和 E_n 值是应用声称不确定度来评价实验室结果的可选方法。如 E_n 值是测量审核结果评定的基本方式，但前提是必须正确评定该实验室对该项测量的不确定度。如果实验室不能正确评定其测量不确定度，则无法使用该方法。能力验证中参加者不确定度的评定可参照 GB/T28043-2019 第 9.8 条。

8.6.5 定性类能力验证统计及评价参见 RB/T186-2017 或 GB/T28043-2019 第 11 章。

8.6.6 能力验证的承担者应将各参加者的检测结果、能力统计量 z 比分数（ z ， zb 和 zw ）、总计统计量（结果总数、中位值、标准化四分位距、稳健变异系数等）归纳并制成各种表格，按照程序来校核数据输入、传送和统计分析的有效性，对所有参加者的检测能力和能力验证的总体情况进行统计分析和评价，确定满意结果、可疑结果、不满意结果、不同结果的参加者名单等。

9 能力验证结果处理

9.0.1 能力验证结果分为满意结果、可疑结果和不满意结果。

9.0.2 对于无故不参加能力验证的参加者、二次能力验证结果为可疑和不满意的参加者，能力验证的组织者应当予以纠正并公布机构名单。

9.0.3 能力验证的参加者私下比对串通能力验证数据、结果，或者提供虚假能力验证数据、结果的，能力验证的组织者应当判定其能力验证结果不满意，并予以公布。

9.0.4 能力验证结果可疑和不满意的参加者，应当及时完成整改，向能力验证的组织者提交整改和验证材料，并经能力验证的组织者确认通过。

9.0.5 能力验证的承担者违反公正性要求，能力验证活动弄虚作假，泄露有关能力验证数据、结果或参加者商业秘密等有关信息的，能力验证的组织者不再委托该其承担能力验证活动。

9.0.6 能力验证的组织者可以通过组织论证、专家评议、监督检查、抽查档案、征求意见等方式，对能力验证活动进行监督。

10 能力验证服务管理平台

10.0.1 能力验证的组织者宜建立符合要求的能力验证服务管理平台，优化能力验证过程，提高能力验证管理信息化水平和工作效率。

10.0.2 能力验证服务管理平台应具备以下功能：

1 组织者可通过平台发布能力验证通知，确定承担者，筛选参加者名单，审核和管理相关信息，能力验证结果发布，接收反馈信息，信息统计查询等。

2 承担者可通过平台提交验证方案、样品均匀性、稳定性检验结果、技术报告，分发样品、能力验证作业指导书，对参加者提交的数据等信息进行审核，数据统计查询等。

3 参加者可通过平台上传能力验证样品接收状态确认单或接收样品的照片、检测结果在线填报、能力验证结果查询等。

4 应能实现对不同用户操作不同功能模块的权限进行设定和管理。

5 应具备数据自动计算的功能，可自动计算出 z 比分数，自动判断参加者的数据是否离群、可疑、满意。

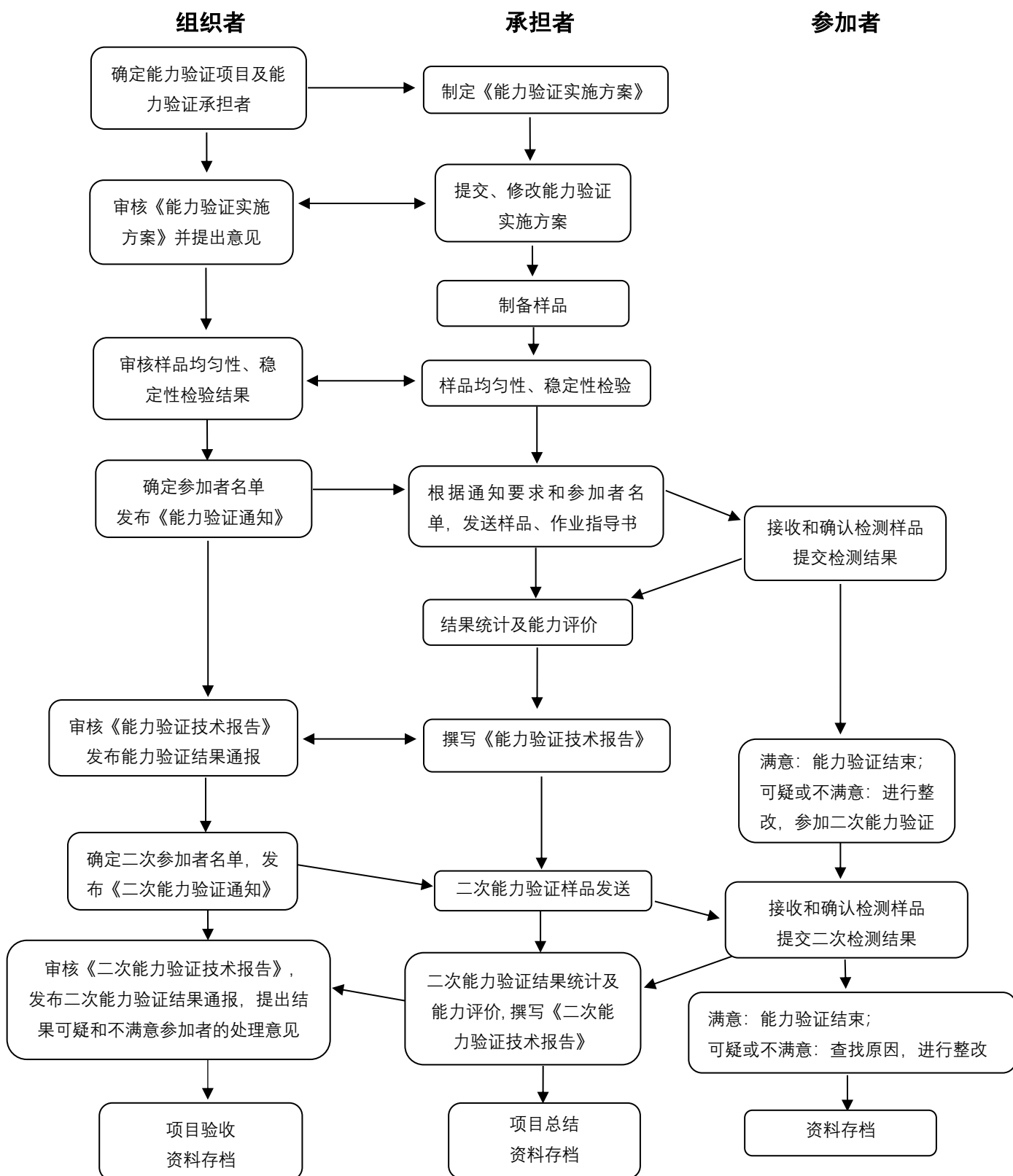
10.0.3 能力验证服务管理平台应确保数据存储与传输安全、可靠。

10.0.4 能力验证服务管理平台系统管理员只能在允许的界面中维护，严禁对系统底层进行操作。

10.0.5 所有进入平台进行操作的事项均应以操作日志形式予以记录。

附录 A

能力验证实施详细流程图



附录B

能力验证实施方案

能力验证项目名称:

能力验证承担单位:

日期:

能力验证项目信息	
能力验证项目名称	
检测参数及对应的检测方法	
承担单位	
能力验证项目组成员信息	
能力验证项目负责人	姓名： 联系电话/手机： 地址： E-mail：
能力验证技术专家	姓名： 联系电话/手机： 地址： 传真： E-mail：
能力验证统计专家	姓名： 联系电话/手机： 地址： 传真： E-mail：
能力验证联络人	姓名： 联系电话/手机： 地址： 传真： E-mail：

能力验证参加者情况：

能力验证样品说明：（包括样品特性、样品拟采用的制备方法，样品均匀性、稳定性检验、样品储存和运输等相关内容）

检测方法说明：（列出本次能力验证项目涉及的产品标准、方法标准，如标准中有检测精度或允差，则应在结果评价时充分考虑）

日程安排：（按组织者能力验证要求，包括方案开始，样品制备、检验、分发，数据收集、处理分析，报告撰写等环节）

数据处理和采用的统计分析、判定原则的说明：（对于一些不适合利用稳健统计方法的特殊计划，可采用专家公议的方式或依据相关行业技术规范进行统计分析和判定，如定性检测类的项目）

计划费用说明和成本预算明细：

其他需要说明的情况：

能力验证实施方案填写说明

B.1 项目信息

项目信息应填写能力验证项目的名称、检测参数及对应的检测方法、承担单位的名称等。

B.2 项目组成员信息

(1) 项目组的成员应包括项目负责人、技术专家、统计专家、联络人信息。其信息应该包含姓名、地址、联系电话或手机、电子邮件等。

(2) 实施方案应当经技术专家、统计专家和项目负责人共同研讨、策划后制定。参与制定方案的人员在能力验证实施和编制报告等方面应具有足够的资格和经验，或能与具有这种能力的人紧密合作。

B.3 参加者的情况

介绍参加者的基本条件、范围和预期数量。

B.4 检测样品说明

检测样品说明应该包括样品特性、样品拟采用的制备方法、样品均匀性稳定性检验、样品储存和运输等相关内容。

(1) 样品的特性：如物理性能（密度、细度、浓度、流动度、强度）等；

(2) 样品拟采用的制备方法：如人工拌合、均匀混合等；

(3) 样品均匀性、稳定性检验。

a. 从样品总体中随机抽取 10 个或 10 个以上的样品用于均匀性检验。若必要，也可以在特性量可能出现差异的部位按一定规律抽取相应数量的检验样品。

b. 对抽取的每个样品，在重复条件下至少测试 2 次。重复测试的样品应分别单独取样。为了减小测量中定向变化的影响（飘移），样品的所有重复测试应按随机次序进行。

c. 均匀性检验中所用的测试方法，其精密度和灵敏度不应低于能力验证预定测试方法的精密度和灵敏度。

d. 特性量的均匀性与取样量有关。均匀性检验所用的取样量不应大于能力验证预定测试方法的取样量。

e. 当检测样品有多个待测特性量时，可从中选择有代表性和对不均匀性敏感的特性量进行均匀性检验。

f. 对检验中出现的异常值，在未查明原因之前，不应随意剔除。

g. 可采用单因子方差分析法对检验中的结果进行统计处理。若样品之间无显著性差异，则表明样品是均匀的。

h. 如果 σ 是某个能力验证中能力评价标准偏差的目标值， S_s 为样品之间不均匀性的标准偏差。若 $S_s \leq 0.3 \sigma$ ，则使用的样品可认为在本能力验证中是均匀的。

(4) 样品储存和运输：如通过特快专递或者快递公司向参加能力验证的实验室发送样

品等。样品在运输和储存过程中，应注意不要损伤外包装，应注意防潮或混入其它杂物。

B.5 检测方法的说明

应当选用参加者日常使用的方法和测试程序，在某些情况下，可以要求参加者采用特定的方法，一般是国家或国际上采纳的标准方法，并已通过适当途径确认。

列出项目计划中涉及到的产品标准、方法标准，如标准中有检测精度或允差，则应在结果评价时充分考虑如标准中有检测精度或允差，则应在结果评价时充分考虑。

B.6 日程安排

(1) 按组织者的要求，样品制备、检验、分发，参加者数据收集、处理，报告撰写等环节。

(2) 各实施步骤应符合组织者的统一时限要求。

B.7 数据处理和采用的统计分析、判定原则的说明

项目承担单位应能对参加者的检测数据进行统计分析，对于一些不适合利用稳健统计方法的特殊计划，可采用专家公议的方式或依据相关行业技术规范进行统计分析和判定，如，定性检测类的项目等。

B.8 计划费用说明和成本预算细目

(1) 一般情况下能力验证计划费用主要包括：样品材料、特殊设备、制样劳务费、搬运劳务费、试验费用、样品包装运输费、统计劳务费等。项目承担单位可根据各自的实际情况制定自己的成本预算细目表。

(2) 山东省建设工程质量检测能力验证计划是非盈利性质的，能力验证项目的承担单位适当收取能力验证相关成本费用。

附录 C

能力验证作业指导书

为保证能力验证的顺利进行及检测操作的一致性，请仔细阅读以下内容并遵照执行。

一、代码

您的检测机构代码为 XXXX，样品代码为 XXXXX。

二、能力验证项目

本次能力验证的检测项目：XXXXXXXX。

三、样品说明

(一)……。

(二)……。

(三)……。

四、检测说明

(一)……。

(二)……。

(三)……。

五、结果提交

(一)……。

(二)……。

(三)……。

六、联系方式

……。

七、其他说明

……。

附录 D

能力验证技术报告

能力验证项目名称:

能力验证承担单位:

日期:

一、能力验证基本信息	
能力验证项目名称	
检测参数及对应的检测标准	
项目实施起止时间	
能力验证承担单位负责人	姓名： 联系电话/手机： 地址： E-mail：
能力验证项目负责人	姓名： 联系电话/手机： 地址： E-mail：
样品制备及均匀性、稳定性试验负责人	姓名： 联系电话/手机： 地址： E-mail：
数据统计分析负责人	姓名： 联系电话/手机： 地址： E-mail：

二、前言（能力验证情况概述等）

三、能力验证样品的均匀性和稳定性检验结果

四、统计分析及能力评价原则

五、统计处理结果及能力评价（包括参加机构名单和评价结果、统计数据）

六、技术分析和建议

七、附件

附件 A 参加者的检测结果和统计处理

A-1 检测结果和统计处理结果

A-2 z 比分数柱状图

附件 B 样品制备和均匀性、稳定性检验

B-1 样品制备报告

B-2 样品均匀性评价报告

B-3 样品稳定性评价报告附录

附件 C 相关文件

C-1 检测能力验证通知文件

C-2 检测能力验证收到样品确认单

C-3 检测能力验证作业指导书

C-4 检测能力验证结果报告单

附件 D 参考文献

附件 E 其他相关材料

能力验证技术报告填写说明

D.1 总则

能力验证技术报告是能力验证项目所有相关信息的汇总，是能力验证的承担者对所承担的能力验证项目的总结。承担者在能力验证实施完成后，及时撰写能力验证技术报告，提交给能力验证的组织者审核。

能力验证项目的技术报告内容根据具体的能力验证项目目的和类型而变化，但都应包含所有参加者的结果分布的数据和每个参加者的能力评价。在编写技术报告时应清晰、全面、简洁。

D.2 报告内容及要求

(1) 能力验证基本信息：应填写能力验证承担者相关人员的信息，包含：承担者的负责人、项目负责人、联系人等，以及样品制备及均匀性、稳定性试验、统计分析等事项的主要负责和参与人员。

(2) 前言：应填写本次能力验证项目的来源，项目的背景和开展的目的，能力验证的项目（参数）和要求，项目实施的效果及意义，以及项目的总体情况，参加者的范围、参加者的数量、参加者的地域分布，样品及参加者采用的编号方法，介绍项目的保密原则，能力验证项目中发现的情况等。

(3) 能力验证样品的均匀性和稳定性检验结果：应进行样品描述，样品的制备细节、包装、标识和发放，样品的均匀性和稳定性检验方法及检验结果（具体数据可列于报告附录中），包括检验方法、取样量、检验次数、检验条件（时间间隔、温度等）和评价的依据和结论等。

(4) 统计分析及相关统计技术的运用说明：介绍能力验证检测结果数据的统计处理方法、结果评价原则及相关统计技术的运用说明，如 z 比分数排列图等。

(5) 统计处理结果及能力评价：总体介绍参加者结果数据统计情况，参加者的代码和检测数据和评价结果（ z 比分数）汇总列表，参加者满意、可疑、离群结果分类统计，可疑结果、不满意参加者代码汇总等。若采用 z 比分数评价检测能力，则按检测指标给出检测结果统计量，包括结果数、中位值、标准化四分位距（NIQR）、稳健的变异系数（CV）、最小值、最大值和极差等，并与均匀性和稳定性检验结果比较，验证统计结果的合理性。给出 z 比分数排列图及数据分布说明、误差的可能来源和能力改进的建议、任何其他建议、意见或一般性评议等。

(6) 技术分析和建议：分析可能导致参加者不满意结果原因分析（如适用），如：不同检测方法/程序对检测结果的影响，不同标准物质对检测结果的影响，检测过程的质量控制因素，参加者对作业指导书的理解程度对结果的影响，原始记录和报告中反映问题等，以

及有关参加者相关检测的改进意见等。总结本次能力验证的实施特点、创新之处和不足，对能力验证参加者、承担者和组织者提出有效的改进建议。

(7) 附件：应包含技术报告需要的支撑材料、其他相关材料等。

附录 E

样品的均匀性和稳定性检验

E.1 基本要求

在实施能力验证时，应保证样品的均匀性和稳定性，以防止样品之间或样品本身的变异导致不满意结果的出现。

E.2 样品均匀性的检验

E.2.1 从能力验证检测项目中选择有代表性和对不均匀性敏感的项目进行均匀性检验，也可选择其他具有相关性且对样品不均匀性敏感的项目进行。

E.2.2 选择执行均匀性检验的实验室和使用的测量方法。测量方法的重复性标准差（ s_r ）应足够小，以便能够检验出任何显著的非均匀性。IUPAC《数据协调议定书》推荐，测量方法的重复性标准差与能力评定标准差的比应小于 0.5（或 $\delta_E/6$ ）。然而该比率要求并非总能达到，这种情况下，能力验证承担者应增加重复测量次数。

E.2.3 在最终完成包装的样品中随机抽取 g 份（ $g \geq 10$ ）进行均匀性检验。若可从前期对相似样品、利用相同程序进行的均匀性检验中得到适合的数据，则可以减少均匀性检验的样本量。

E.2.4 为一轮能力验证制备并包装样品，应保证为该轮能力验证的参加者以及均匀性检验提供充足的样品。

E.2.5 从一个样品中制备 $m \geq 2$ 个测试份量，选择适合的方法将测试份量间差异降到最小。将 $g \times m$ 份测试份量随机排序，得到每份的测量结果，整个测量在重复性条件下完成。若不能进行重复测量，比如破坏性测试，则测量结果的标准差可用作 s_s 。这种情况下，使用具有足够低的重复性标准差 s_r 的方法是非常重要的。

E.2.6 若 σ_{pt} 事先未知，比如 σ_{pt} 是参加者测量结果的稳健标准差，能力验证承担者宜选用判定充分均匀性的其他标准。该程序可能包括：

- a) 采用方差分析检验法（ $\alpha = 0.05$ ）对检验中的结果进行统计处理，若样品之间无显著性差异，则表明样品是均匀的；
- b) 使用以往能力验证获得的信息估计 σ_{pt} ；
- c) 使用精密度实验获得的数据（比如 GB/T6379.2 中描述的再现性标准差）；
- d) 使用不具备充分均匀性的能力验证样品，接受其风险，并在计算出公议值之后，对评估标准进行核查。

E.2.7 如果 σ_{pt} 是某个能力验证中能力评价标准偏差的目标值， s_s 为样品之间不均匀性的标准偏差，可采用 $s_s \leq 0.3 \sigma_{pt}$ 准则。即若 $s_s \leq 0.3 \sigma_{pt}$ ，则使用的样品可认为在本能力验证中

是均匀的。

E.2.8 如果不符合充分均匀性的标准，能力验证承担者应考虑采取如下措施之一：

a) 将样本间标准差计入能力评定标准差中，利用下式计算 σ' 。这需要向参加者详细说明。

$$\sigma' = \sqrt{\sigma_{pt}^2 + s_s^2}$$

b) 在指定值的不确定度计算中使用 s_s ，并在能力评定时使用 z' 或 δ'_E 。

c) 若 σ_{pt} 是参加者结果的稳健标准差，则 σ_{pt} 包含能力验证样品之间的非均匀性，因此，能力验证样品均匀性的标准可谨慎放宽。

若 a) ~ c) 均不适用，则能力验证样品作废，在满足均匀性检验后重新制备。

E.2.9 方差分析 F 检验

抽取 i 个样品 ($i=1, 2, \dots, m$)，每个样品在重复条件下测试 j 次 ($j=1, 2, \dots, n$)。

每个样品的测试平均值
$$\bar{x}_i = \sum_{j=1}^n x_{ij}/n_i$$

全部样品测试的总平均值
$$\bar{x} = \sum_{i=1}^m \bar{x}_i/m$$

测试总次数
$$N = \sum_{i=1}^m n_i$$

样品间平方和
$$SS_1 = \sum_{i=1}^m n_i(\bar{x}_i - \bar{x})^2 \quad \text{均方 } MS_1 = \frac{SS_1}{f_1}$$

样品内平方和
$$SS_2 = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \quad \text{均方 } MS_2 = \frac{SS_2}{f_2}$$

自由度
$$f_1 = m - 1$$

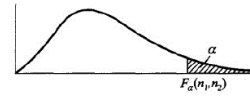
$$f_2 = N - m$$

统计量
$$F = \frac{MS_1}{MS_2}$$

若 $F <$ 自由度为 (f_1, f_2) 及给定显著性水平 α (通 $\alpha=0.05$) 的临界值 $F_\alpha(f_1, f_2)$ (查表 E1 F 分布表), 则表明样品内和样品间无显著性差异, 样品是均匀的。

表 E1 F 分布表

$$P\{F(f_1, f_2) > F_\alpha(f_1, f_2)\} = \alpha \quad \alpha = 0.05$$



$f_2 \backslash f_1$	8	9	10	11	12	13	14	15	16	17	18	19	20
9	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94
10	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77
11	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65
12	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54
13	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46
14	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39
15	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33
16	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28
17	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23
18	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19
19	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16
20	2.45	2.39	2.35	2.31	2.28	2.25	2.23	2.20	2.18	2.17	2.15	2.14	2.12

(摘自 GB 4086.4-1983 统计分布数值表 F 分布)

注: 在 excel 工作表任意单元格输入公式 “=FINV(0.05, f_1 , f_2)” 后回车亦可计算得到, 其中 f_1 , f_2 为计算得到的自由度。

E.2.10 $S_s \leq 0.3 \sigma_{pt}$ 准则

从制备的样品中随机抽取 i 个样品 ($i=1, 2, \dots, m$), 每个样品在重复条件下测试 j 次 ($j=1, 2, \dots, n$)。

按 E.2.9 计算均方 MS_1 、 MS_2 。

若每个样品的重复测试次数均为 n 次。按下式计算样品之间的不均匀性标准偏差 S_s :

$$S_s = \sqrt{(MS_1 - MS_2)/n}$$

式中: MS_1 —— 样品间均方;

MS_2 —— 样品内均方;

n —— 测量次数。

若 $S_s \leq 0.3 \sigma_{pt}$, 则使用的样品可认为在本能力验证中是均匀的。式中 σ_{pt} 是能力验证中能力评价标准偏差的目标值。

E.3 样品的稳定性检验

E.3.1 对于某些性质较不稳定的检测样品, 运输和时间对检测的特性量可能会产生影响。因此需要进行稳定性检验。当检测样品有多个待测特性量时, 应选择容易发生变化和有代表性的特性量进行稳定性检验。

E.3.2 稳定性检验的测试方法应是精密和灵敏的, 并且具有很好的复现性。

E.3.3 稳定性检验的样品应从包装单元中随机选择数量为 $2g$ 的能力验证样品，其中 $g \geq 2$ ，抽取的样品数应具有足够的代表性。

E.3.4 应在样品发送之前和所有参加者完成能力验证样品检测后进行稳定性检验。在向参加者分发能力验证样品的日期之前，对数量为 g 的能力验证样品以完全随机的顺序进行重复测量；在与参加者预期存储条件类似的条件下，保存数量为 g 的剩余能力验证样品；在参加者返回测量结果的截止日期之后，使用相同实验室、相同测量方法和相同数量的重复测量，在合理的期限内尽快对数量为 g 的剩余能力验证样品进行测量，且所有的测量应采用随机顺序。

E.3.5 能力验证承担者应至少在执行能力验证的早期阶段，检验运输条件对能力验证样品的影响。在可能的情况下，应将能力验证承担者保留的能力验证样品与经过运输并返回的能力验证样品进行比较。也可使用根据暴露于合理可预见的运输条件时研究结果。

进行能力评定时应考虑任何已知的运输影响。指定值的不确定度宜包含运输条件所导致的不确定度的任何显著增加。

如果运输稳定性检验涉及两组能力验证样品测量结果的比较，即一组样品经过实际运输，另一组样品未经过实际运输，则运输条件的稳定性准则与 E.3.6 的规定相同。

E.3.6 $|\bar{x} - \bar{y}| \leq 0.3 \sigma_{pt}$ 准则：

若 $|\bar{x} - \bar{y}| \leq 0.3 \sigma_{pt}$ 成立，则认为被检的样品是稳定的。

式中： \bar{x} ——均匀性检验的总平均值；

\bar{y} ——稳定性检验时，对随机抽出样品的测量平均值。

注：抽样数 ≥ 3 。对每个抽取的样品重复测试 2 次，每次分别单独取样。测量方法与均匀性检验用的测量方法相同。

σ_{pt} 该能力验证的能力评价标准偏差目标值。

E.3.7 t 检验法

E.3.7.1 一系列测量的平均值与标准值/参考值的比较

按下式计算 t 值：

$$t = \frac{|\bar{x} - \mu| \sqrt{n}}{S}$$

式中： \bar{x} —— n 次测量的平均值；

μ ——标准值/参考值；

n ——测量次数；

S —— n 次测量结果的标准偏差。

注：为了保证平均值和标准偏差的准确度， $n \geq 6$ 。

若 $t <$ 显著性水平 α （通常 $\alpha = 0.05$ ）自由度为 $n - 1$ 的临界值 $t_{\alpha(n-1)}$ （查表 E2 t 分布表），则平均值与标准值/参考值之间无显著性差异。

E.3.7.2 二个平均值之间的一致性

按下式计算 t 值:

$$t = \frac{|\bar{x}_2 - \bar{x}_1|}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 \cdot n_2}}}$$

式中: \bar{x}_1 —— 第一次检验测量数据的平均值;

\bar{x}_2 —— 第二次检验测量数据的平均值;

s_1 —— 第一次检验测量数据的标准偏差;

s_2 —— 第二次检验测量数据的标准偏差;

n_1 —— 第一次检验测量的测量次数;

n_2 —— 第二次检验测量的测量次数。

注: 为了保证平均值和标准偏差的准确度, n_1 和 n_2 均 ≥ 6 。

若 $t <$ 显著性水平 α (通常 $\alpha = 0.05$) 自由度为 $n_1 + n_2 - 2$ 的临界值 $t_{\alpha(n_1 + n_2 - 1)}$ (查表 E2 t 分布表), 则二个平均值之间无显著性差异。

表 E2 t 分布表

自由度	10	11	12	13	14	15	16	17	18	19
临界值	2.2281	2.2010	2.1788	2.1604	2.1448	2.1315	2.1199	2.1098	2.1009	2.0930
自由度	20	21	22	23	24	25	26	27	28	29
临界值	2.0860	2.0796	2.0739	2.0687	2.0639	2.0595	2.0555	2.0518	2.0484	2.0452
自由度	30	31	32	33	34	35	36	37	38	39
临界值	2.0423	2.0395	2.0369	2.0345	2.0322	2.0301	2.0281	2.0262	2.0244	2.0227

(摘自 GB 4086.4-1983 统计分布数值表 F 分布)

注: 在 excel 工作表任意单元格输入公式 “=TINV(0.05, $n_1 + n_2 - 2$)” 后回车亦可计算得到, 其中 $n_1 + n_2 - 2$ 为自由度。

E.4 样品的均匀性检验结果示例

表 E3 ***液限含水率均匀性检验分析结果 (%)

样品编号 (i)	测试次数 (j)	
	1	2
1	25.4	25.2
2	24.5	25.1
3	24.6	24.9
4	25.2	25
5	25.0	24.8
6	25.1	25.5
7	24.7	24.9
8	25.0	25.3
9	25.5	24.6
10	25.0	25.4
总平均值	25	

表 E4 ***方差分析结果

方差来源	平方和	自由度	均方	F 值
样品间	0.8	9	0.089	0.99
样品内	0.9	10	0.09	

查表 E1 得到 F 临界值 $F_{0.05}(9, 10) = 3.02$ 。计算的 F 值为: 0.99, 该值小于 F 临界值, 这表明在 0.05 显著性水平时, 样品是均匀的。

附录 F

检测能力验证样品确认单

能力验证项目名称			
承担机构			
发送人		电话/传真	
运输单据号		发送状态	完好 <input type="checkbox"/> 不完好 <input type="checkbox"/>
以下由样品接收机构填写			
接收机构名称			
接收机构代码		样品代码	
联系地址			
接收人		电话/传真	
收到样品日期			
<p>接收时，被测样品状态是否良好： 是<input type="checkbox"/> 否 <input type="checkbox"/></p> <p>如需要，对接收状态的详细说明：</p> <p style="text-align: center;">接收人签名：</p> <p style="text-align: center;">年 月 日</p>			
备注：			

附录 G

结果统计分析方法

G.1 概述

检测能力验证结果统计分析方法通常可以采用经典方法，用格拉布斯（Grubbs）准则等统计方法剔除离群值后计算平均值和标准差，以平均值和标准差作为指定值和能力评定标准差；也可采用稳健统计方法，稳健统计方法不需要用统计方法剔除离群值。稳健统计方法主要是一种在实验数据处理中，不会被一些异常值干扰的统计学方法。由于参加者结果通常不是呈正态分布而是偏态分布，因而，人们运用了稳健统计技术，规定了平均值估计值和标准差估计值，利用它们稳健的特性在能力验证过程中评价所有参加实验室检测数据结果的准确性，在不将极端结果从数据组中剔除的情况下使其的影响减至最小。所谓稳健就是即使结果数据中存在异常值（离群值）的影响，仍能给出较好的结果。

本附录内容主要包括两种常用且较为简便的能力验证结果稳健统计方法：算法 A、标准化四分位距法，其它方法可参照 GB/T28043-2019 附录 C。

此外本附录还对参加者较少时的统计方法、混凝土结构实体强度现场能力验证统计方法进行了阐述。

G.2 算法 A

算法 A 通过迭代方法转化原始数据，为近似正态分布提供均值和标准偏差的替代计算方法。迭代算法 A 较四分位法更能真实地反映实验室之间检测结果的变异性。需要进行计算的数据、计算方法及结果评定如下：

G.2.1 总体平均值和标准差的稳健值

假设参加者（实验室）提交的 p 个数据按递增顺序排列表示为： $x_1, x_2, \dots, x_i, \dots, x_p$ 。这些数据的稳健平均值和稳健标准差记为 x^* 和 s^* 。先计算 p 个数据的中位值作为初始稳健平均值（ x^* ），计算其绝对中位差作为初始稳健标准差（ s^* ）。

计算 x^* 和 s^* 的初始值如下（med表示中位值）：

$$x^* = \text{med}\{x_i\}(i = 1, 2, \dots, p) \quad \dots\dots\dots (G.1)$$

$$s^* = 1.483 * \text{med}\{|x_i - x^*|\}(i = 1, 2, \dots, p) \quad \dots\dots\dots (G.2)$$

根据以下步骤更新 x^* 和 s^* 的值。计算：

$$\delta = 1.5s^* \quad \dots\dots\dots (G.3)$$

对每个 $x_i (i=1, \dots, p)$ ，计算：

$$x_i^* = \begin{cases} x^* - \delta & (\text{当 } x_i < x^* - \delta) \\ x^* + \delta & (\text{当 } x_i > x^* + \delta) \\ x_i & (\text{在其他情况下}) \end{cases} \dots\dots\dots (G.4)$$

再由下式计算 x^* 和 s^* 的新的取值:

$$x^* = \frac{\sum_{i=1}^p x_i^*}{p} \dots\dots\dots (G.5)$$

$$s^* = 1.134 \sqrt{\frac{\sum_{i=1}^p (x_i^* - x^*)^2}{p-1}} \dots\dots\dots (G.6)$$

稳健估计值 x^* 和 s^* 可由迭代计算得出, 例如用新取值数据更新 x^* 和 s^* , 直至过程收敛。当稳健平均值和稳健标准差的第三位有效数字在连续两次迭代中不再变化时, 即可认为过程是收敛的。这是一种可用计算机编程 (或 excel) 实现的简单方法, 示例见 G.2.6。

G.2.2 指定值及其不确定度

可由某轮能力验证计划参加者 (实验室) 报告的结果计算得出的稳健平均值 (x^*) 作为能力验证样品的指定值 x_{pt} ;

当利用算法 A 的稳健统计方法确定指定值时, 指定值 x_{pt} 的标准不确定度可按下式计算:

$$u(x_{pt}) = 1.25 \times \frac{s^*}{\sqrt{p}} \dots\dots\dots (G.7)$$

式中: $u(x_{pt})$ 为稳健平均值的标准不确定度;

s^* 是结果的稳健标准差;

p 为参加者 (实验室) 数。

注1: 指定值和稳健标准差由参加者 (实验室) 结果获得, 可假定指定值的不确定度包含不均匀性、运输和不稳定性对不确定度的影响。

注2: 基于正态分布的大量数据的中位值的标准差, 或是中位值作为平均值估计值的效率, 确定校正系数为1.25。较复杂的稳健统计方法的效率远大于中位值效率, 因此校正系数小于1.25, 之所以选用该系数, 是因为能力验证结果一般不严格服从正态分布, 且包含未知比例的源于不同分布的结果 (“污染结果”)。就可能存在的 “污染结果” 而言, 1.25是一个保守 (高) 估计值, 可依据经验和所使用的稳健方法选择较小的校正系数或不同的公式。

G.2.3 能力评定标准差

能力评定标准差 σ_{pt} 可由同一轮能力验证参加者 (实验室) 结果计算得出, 使用该方法计算能力统计量时, 适宜使用 z 值, 通常使用稳健标准差 (s^*) 作为所有参加者 (实验室) 报告结果的稳健标准差 (σ_{pt})。

G.2.4 z 值

能力验证结果 x_i 的 z 值的计算公式如下:

$$z_i = \frac{x_i - x_{pt}}{\sigma_{pt}} \dots\dots\dots (G.8)$$

式中： x_i 表示一轮能力验证中各参加者（实验室） i 的测量结果（或重复测试结果的平均值）； σ_{pt} 即当稳健平均值（ x^* ）作为能力验证样品的指定值时（ x^* ）的值。

G.2.5 结果评定

z 值的评定标准为：

当 $|z| \leq 2.0$ 时，表示满意结果；

$2.0 \leq |z| \leq 3.0$ 时，表示可疑结果；

$|z| \geq 3.0$ 时，表示不满意结果。

注1：某些要求严格的应用中， $|z| > 2.0$ 视为不满意结果。

注2：将 z 值的限值确定为 2.0 或 3.0 的原因如下：若正确开展的测量得到的结果（如有必要需对结果进行转化）服从正态分布，其均值为 x_{pt} ，标准差为 σ_{pt} 。那么 z 值也将呈正态分布，其均值为 0，标准差为 1.0。在此条件下， z 值落在 $-3.0 \leq z \leq 3.0$ 的区间之外的概率只有 0.3%，而在 $-2.0 \leq z \leq 2.0$ 的区间之外的概率则有 5%。因为不满意发生概率很低，当没有真正的问题存在时很难出现不满意，所以若出现了不满意则认为有异常情况出现是合理的。

注3：如果实际的实验室间差异小于 σ_{pt} ，则误判的概率会降低。

需根据所报告 z 值的有效数字的位数来确定结果、指定值和能力评定标准差的有效数字。对于 z 值，通常保留小数点后两位数字。当参加者数量很大且参加者的结果的标准差用作 σ_{pt} 时，建议承担者检查参加者结果或 z 值分布的正态性（可参考 GB/T28043-2019 第 10 章）。

G.2.6 算法 A 计算示例（excel）

稳健估计值 x^* 和 s^* 的迭代计算是一种可用 excel（或计算机编程）实现的简单方法，下面介绍一下如何使用 excel 公式函数及公式进行稳健估计值 x^* 和 s^* 的迭代计算及能力验证结果评定。

将某次能力验证中混凝土试块的抗压强度项目的各参加者试验结果录入到 excel 中，按从小到大顺序排列： $x_1, x_2, \dots, x_i, \dots, x_p$ （ $p = 43$ ），其中位值或可称为初始稳健平均值为 46.4MPa。计算各参加者试验结果与初始稳健平均值（ x_0^* ）之差的绝对值 $|x_i - x_0^*|$ 组成新数列，该数列的中位值 1.7 MPa 按公式 G.4 计算初始稳健标准差为 2.52；修正后为 3.78，按公式 G.4-G.6 多次修正迭代，直至稳健平均值与稳健标准偏差稳定至 46.2 和 4.07。

如指定值为参加者确定的公议值确定，那么此稳健平均值 46.2 即为能力验证结果的指定值。

由得到的稳健平均值与稳健标准偏差可计算出能力评定统计量 z 比分数。

具体数据及能力验证结果评定如下表所示：

表 G1 算法 A 能力验证结果统计表

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	实验室代码	$ x_i - x_0 $	x_i	迭代1	迭代2	迭代3	迭代4	迭代5	迭代6	迭代7	迭代8	...	Z比分数值	评定结论
7	034	4.6	41.8	42.61835	42.44903	42.34269	42.27523	42.23216	42.20456	42.18683	42.17664		-1.09116	满意
8	035	4.2	42.2	42.61835	42.44903	42.34269	42.27523	42.23216	42.20456	42.2	42.2		-0.99288	满意
9	021	3.3	43.1	43.1	43.1	43.1	43.1	43.1	43.1	43.1	43.1		-0.77174	满意
10	039	2.6	43.8	43.8	43.8	43.8	43.8	43.8	43.8	43.8	43.8		-0.59975	满意
11	040	2.3	44.1	44.1	44.1	44.1	44.1	44.1	44.1	44.1	44.1		-0.52604	满意
12	018	2.1	44.3	44.3	44.3	44.3	44.3	44.3	44.3	44.3	44.3		-0.4769	满意
13	029	1	45.4	45.4	45.4	45.4	45.4	45.4	45.4	45.4	45.4		-0.20662	满意
14	004	0.8	45.6	45.6	45.6	45.6	45.6	45.6	45.6	45.6	45.6		-0.15748	满意
15	024	0.8	45.6	45.6	45.6	45.6	45.6	45.6	45.6	45.6	45.6		-0.15748	满意
16	046	0.3	46.1	46.1	46.1	46.1	46.1	46.1	46.1	46.1	46.1		-0.03463	满意
17	013	0.2	46.2	46.2	46.2	46.2	46.2	46.2	46.2	46.2	46.2		-0.01006	满意
18	007	0.1	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3		0.014515	满意
19	020	0.1	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3		0.014515	满意
20	030	0.1	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3		0.014515	满意
21	037	0.1	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3		0.014515	满意
22	038	0.1	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3		0.014515	满意
23	009	0	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4		0.039085	满意
24	010	0	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4		0.039085	满意
25	027	0	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4		0.039085	满意
26	031	0	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4	46.4		0.039085	满意
27	017	0.8	47.2	47.2	47.2	47.2	47.2	47.2	47.2	47.2	47.2		0.23565	满意
28	044	0.8	47.2	47.2	47.2	47.2	47.2	47.2	47.2	47.2	47.2		0.23565	满意
29	001	1.1	47.5	47.5	47.5	47.5	47.5	47.5	47.5	47.5	47.5		0.309361	满意
30	008	1.2	47.6	47.6	47.6	47.6	47.6	47.6	47.6	47.6	47.6		0.333932	满意
31	045	1.2	47.6	47.6	47.6	47.6	47.6	47.6	47.6	47.6	47.6		0.333932	满意
32	006	1.4	47.8	47.8	47.8	47.8	47.8	47.8	47.8	47.8	47.8		0.383073	满意
33	016	1.7	48.1	48.1	48.1	48.1	48.1	48.1	48.1	48.1	48.1		0.456784	满意
34	042	1.7	48.1	48.1	48.1	48.1	48.1	48.1	48.1	48.1	48.1		0.456784	满意
35	015	1.9	48.3	48.3	48.3	48.3	48.3	48.3	48.3	48.3	48.3		0.505925	满意
36	043	2	48.4	48.4	48.4	48.4	48.4	48.4	48.4	48.4	48.4		0.530496	满意
37	019	2.2	48.6	48.6	48.6	48.6	48.6	48.6	48.6	48.6	48.6		0.579637	满意
38	023	2.4	48.8	48.8	48.8	48.8	48.8	48.8	48.8	48.8	48.8		0.628778	满意
39	036	2.7	49.1	49.1	49.1	49.1	49.1	49.1	49.1	49.1	49.1		0.70249	满意
40	025	2.8	49.2	49.2	49.2	49.2	49.2	49.2	49.2	49.2	49.2		0.72706	满意
41	032	2.9	49.3	49.3	49.3	49.3	49.3	49.3	49.3	49.3	49.3		0.751631	满意
42	003	3	49.4	49.4	49.4	49.4	49.4	49.4	49.4	49.4	49.4		0.776201	满意
43	002	3.2	49.6	49.6	49.6	49.6	49.6	49.6	49.6	49.6	49.6		0.825342	满意
44	022	3.6	50	50	50	50	50	50	50	50	50		0.923625	满意
45	\bar{x}^*	46.4		46.31229	46.28473	46.26741	46.25643	46.24942	46.24493	46.24235	46.24093			
46	s^*	2.5211		2.575509	2.628022	2.661459	2.682849	2.696574	2.705396	2.710473	2.713277			
47	$\bar{\sigma}$	3.78165		3.863264	3.942034	3.992188	4.024274	4.04486	4.058095	4.065709	4.069915			
48	$x^* - \bar{\sigma}$	42.61835		42.44903	42.34269	42.27523	42.23216	42.20456	42.18683	42.17664	42.17101			
49	$x^* + \bar{\sigma}$	50.18165		50.17555	50.22676	50.2596	50.28071	50.29428	50.30302	50.30806	50.31084			
50	s_0^*	1.7												
51	x_0^*	46.4												
52	稳健平均值 x_{ST}									46.2	46.2			
53	稳健标准差 σ_{ST}									4.07	4.07			

上表各单元格所用excel函数及公式如下：

A2: A44为各参加者（实验室）代码；单元格C2: C44录入43个结果值；将A1: C44选中，在“数据”菜单中选择“排序”，排序依据选择以列C单元格值升序排列。然后依次在下列单元格中输入公式：

C51=MEDIAN(C2:C44); B2=ABS(C2-C\$51)，选中B2下拉至B44；

C50=MEDIAN(B2:B44)；

C45=C51； C46=1.483*C50； C47=1.5*C46； C48=C45-C47； C49=C45+C47；

D2=IF(\$C2<C\$48,C\$48,IF(\$C2>C\$49,C\$49,\$C2))，选中D2下拉至D44；

D45=AVERAGE(D2:D44)； D46=1.134*STDEV(D2:D44)； D47=1.5*D46；

D48=D45-D47；

D49=D45+D47；

选中D1: D49向右拖至任意一列，即可完成多次迭代运算（本例为8次）；

M2 =(C2-\$K\$52)/\$K\$53，选中M2下拉至M44；

$N2 = IF(ABS(M2) \leq 2, "满意", IF(AND(ABS(M2) > 2, ABS(M2) < 3), "可疑", "不满意"))$ ，然后选中N2下拉至N44即可完成能力评定。

G.2.7 z比分数柱状图示例

根据表 G1 算法 A 能力验证结果统计表中的 z 比分数数据作图，得到 z 比分数柱状图如下：

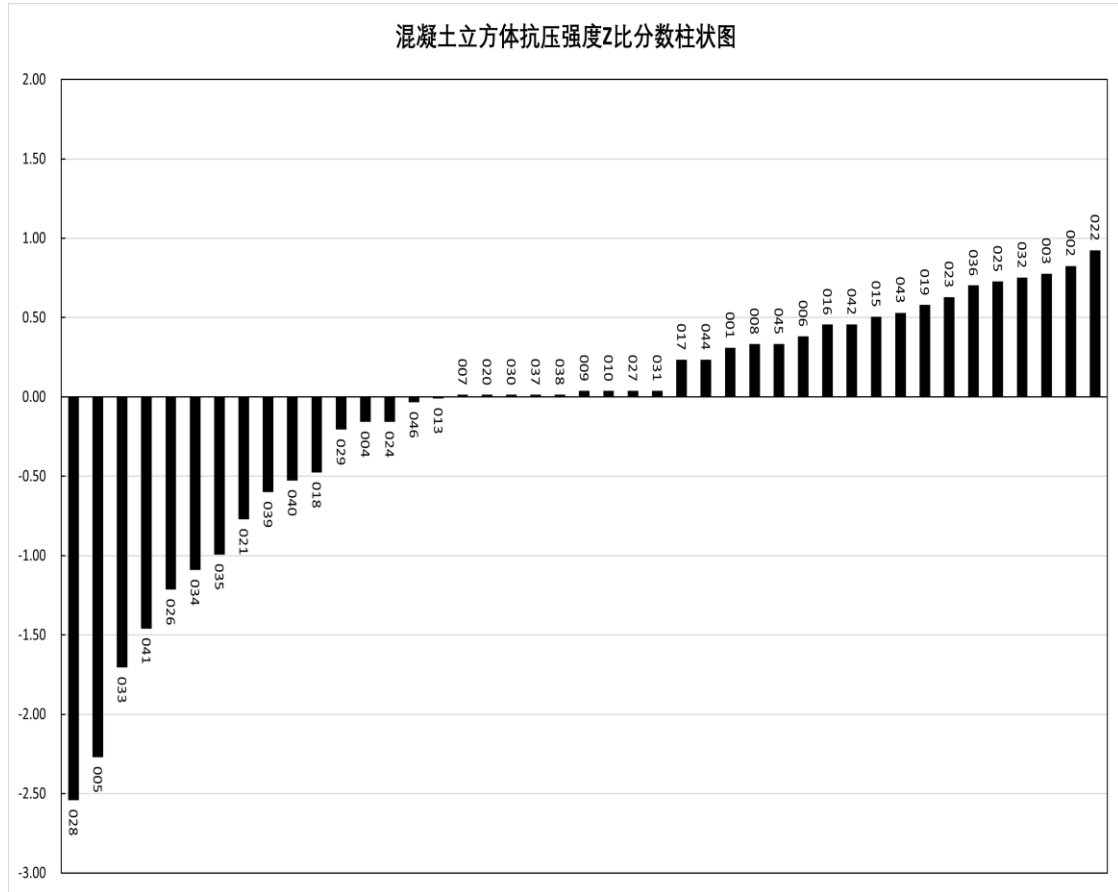


图 G1 z 比分数柱状图

G.3 标准化四分位距法

四分位数是统计学中分位数的一种，即把所有数值由小到大排列并分成四等份，处于三个分割点位置的数值就是四分位数。第一四分位数（Q1），又称“较小四分位数”，等于该样本中所有数值由小到大排列后第 25% 的数字。第二四分位数（Q2），又称“中位数”，等于该样本中所有数值由小到大排列后第 50% 的数字。第三四分位数（Q3），又称“较大四分位数”，等于该样本中所有数值由小到大排列后第 75% 的数字。第三四分位数与第一四分位数的差距又称四分位距（IQR）。标准化四分位距法需要进行计算的数据、计算方法及结果评定如下：

G.3.1 中位值

中位值是对称分布总体平均值的一种简单估计，该方法对离群值不敏感，可用 $med(x)$

表示中位值。假设参加者提交的 p 个数据按递增顺序表示为： $x_1, x_2, \dots, x_i, \dots, x_p$ 。当 p 为奇数时，中位值为第 $x_{(\frac{p+1}{2})}$ 位的数值；当 p 为偶数时，中位值为第 $x_{(\frac{p}{2})}$ 位和第 $x_{(\frac{p}{2}+1)}$ 位数值值的平均值。可用下式表示：

$$\text{med}(x) = \begin{cases} x_{(\frac{p+1}{2})} & (\text{当 } p \text{ 为奇数时}) \\ \frac{x_{(\frac{p}{2})} + x_{(\frac{p}{2}+1)}}{2} & (\text{当 } p \text{ 为偶数时}) \end{cases} \dots\dots\dots (\text{G.9})$$

G.3.2 四分位值和四分位距 (IQR)

上四分位值 Q_3 又称高四分位值，它是一组按顺序排列的数值，尽可能有 1/4 的数值高于该值，即有 3/4 (75%) 的数值低于该值；下四分位值 Q_1 ，又称低四分位值，它是在一组顺序排列的数值中，尽可能有 1/4 (25%) 的数值低于该值，即有 3/4 的数值高于该值。当 p 满足 $p=9+4n$ ($n=0,1,2,\dots$) 时， Q_3 和 Q_1 值在数列中的位置为整数，其它情况下 Q_3 和 Q_1 之值在数列中的位置需通过内插法求得，故在安排能力验证时，应尽可能按 $p=9+4n$ 加以考虑，这会给数据处理带来很大的方便。

在实验结果从小到大排列的数值中，中位值 $\text{med}(x)$ 的位置是：

$$Q_{\text{中}} = \frac{1}{2}(N + 1) \dots\dots\dots (\text{G.10})$$

下四分位值 Q_1 的位置是：

$$Q'_1 = \frac{1}{2}(Q_{\text{中}} + 1) \dots\dots\dots (\text{G.11})$$

上四分位值 Q_3 的位置是：

$$Q'_3 = Q_{\text{中}} + Q'_1 - 1 \dots\dots\dots (\text{G.12})$$

例 1：一组数据从小至大排列为：1.0,1.3,2.0,4.2,5.0,6.3,6.5,7.0,7.2,8.0,8.1,8.6,9.3, 9.5,10.5,10.6,12.0 ($p=17$)

中位值 x_m 的位置是 $\frac{1}{2} \times (17 + 1) = 9$ ，即 $\text{med}(x) = x_9 = 7.2$

下四分位值 Q_1 的位置是 $\frac{1}{2} \times (9 + 1) = 5$ ，即 $Q_1 = x_5 = 5.0$

上四分位值 Q_3 的位置是 $9 + 5 - 1 = 13$ ，即 $Q_3 = x_{13} = 9.3$

例 2：一组数据从小至大排列为：1.0,1.3,2.0,4.2,5.0,6.2,6.5,7.0,7.2,8.0 ($p=10$)

中位值 x_m 的位置是：

$$\frac{1}{2} \times (10 + 1) = 5.5 \text{ 即 } \text{med}(x) = x_{5.5} = \frac{1}{2} \times (x_5 + x_6) = \frac{1}{2} \times (5.0 + 6.2) = 5.6$$

下四分位值 Q_1 的位置是：

$$\frac{1}{2} \times (5.5 + 1) = 3.25 \text{ 即 } Q_1 = x_{3.25} = x_3 + 0.25(x_4 - x_3) = 2.55$$

上四分位值 Q_3 的位置是：

$$5.5 + 3.25 - 1 = 7.75 \text{ 即 } Q_3 = x_{7.75} = x_7 + 0.75(x_8 - x_7) = 7.15$$

上四分位值 Q_3 与下四分位值 Q_1 之间的差称为四分位距 (IQR), 即:

$$IQR = Q_3 - Q_1 \dots\dots\dots (G.13)$$

G.3.3 标准化四分位距 nIQR

标准化四分位距等于四分位距 (IQR) 乘以因子 0.7413。

可将参加者结果按递增顺序排列, 计算第 75 百分位 (或第三个四分位) Q_3 和第 25 百分位 (或第一个四分位) Q_1 参加者结果的差值, 然后乘以系数 0.7413 即可得到标准化四分位距。可按下式计算得到:

$$nIQR = 0.7413(Q_3 - Q_1) \dots\dots\dots (G.14)$$

式中:

Q_1 —— x_i ($i=1, \dots, p$) 的 25%分位数;

Q_3 —— x_i ($i=1, \dots, p$) 的 75%分位数。

G.3.4 稳健 z 比分数、实验室间 z 比分数 (zb)、实验室内 z 比分数 (zw)

为统计和评定参加能力验证的试验结果, 采用根据稳健总统计量 (如中位值 $\text{med}(x)$, nIQR 等) 计算的 z 比分数为稳健 z 比分数, 简称 z 比分数, 即:

$$z = \frac{\text{结果} - \text{中位值}}{nIQR} \dots\dots\dots (G.15)$$

当采用分割水平样品 (样品对) 设计能力验证时, 将计算两个 z 比分数: 实验室间 z 比分数 (zb) 和实验室内 z 比分数 (zw)。

假设样品对是 A 和 B 两个样品, 结果对的标准化和为 S , 标准化差为 D , 则:

$$S = \frac{(A + B)}{\sqrt{2}} \dots\dots\dots (G.16)$$

$$D = \frac{(A - B)}{\sqrt{2}} \text{ (若 } A \text{ 样的中位值} > B \text{ 样的中位值)} \dots\dots\dots (G.17)$$

或

$$D = \frac{(B - A)}{\sqrt{2}} \text{ (若 } A \text{ 样的中位值} < B \text{ 样的中位值)} \dots\dots\dots (G.18)$$

在能力验证中, 通过计算每一个参加实验室的标准化和 S 及标准化差 D , 可以得出 S 与 D 的中位值和标准 IQR, 一组 S 数据的中位值记为中位值 (S), 其标准化四分位距为标准 IQR (S)。随后计算实验室间 z 比分数 (zb) 和实验室内 z 比分数 (zw), 计算公式为:

$$zb = \frac{S - \text{中位值}(S)}{nIQR(S)} \dots\dots\dots (G.19)$$

式中 S 的稳健 z 比分数 zb 的大小代表某试验室的 S (标准化和) 与中位值 (S) 的偏离程度, 符号“+”与“-”代表 S 与中位值 (S) 的偏离方向, “+”为偏大, “-”为偏小。此值反映了实验室测定结果的系统误差。

$$zw = \frac{D - \text{中位值}(D)}{nIQR(D)} \dots\dots\dots (G.20)$$

式中 D 的稳健 z 比分数 z_w 的大小和符号代表某实验室的标准化差 D 与中位值(D)的偏离程度和方向。此值反映了实验室测定结果的随机误差。

如果要求每个参加者对每个验证样的每个测试项目重复测试 2 次或以上时, 应利用多次测试结果的平均值计算 z_b 或 z_w 。

当能力验证只能获得单一验证样的结果, 特别是需通过破坏性试验才能获得结果, 并且具有一定分散性的项目参数(如金属材料的吸收冲击功), 在此类能力验证中, 每个参加者收到 3 个(至少 2 个)验证样, 每个参加者按照能力验证对有效位数和数字修约的要求报出每个验证样的结果或其平均值, 此时应将平均值作为检测结果, 计算出每个实验室的能力统计量 z 比分数。

z 比分数 z_b 或 z_w 值越小, 偏离程度越小, 实验室能力越强; 反之, 此值越大, 偏离程度越大, 实验室能力越差。 z 比分数的评定标准为:

当 $|z| \leq 2.0$ 时, 表示满意结果;

$2.0 \leq |z| \leq 3.0$ 时, 表示可疑结果;

$|z| \geq 3.0$ 时, 表示不满意结果。

正的 z 比分数表明结果偏高, 负的 z 比分数表明结果偏低。这时, 实验室必须认真进行整改, 纠正偏离。

当 $2 < |z| < 3$ 时, 实验数据为存在问题的结果, 是统计复查值, 相应的实验室应该进行复查。

G.3.5 稳健变异系数

以标准化四分位距除以中位值, 并以百分数来表示, 即:

$$C_v = \frac{nIQR}{\text{中位值}} \times 100\% \dots\dots\dots (G.21)$$

此参数可以比较不同样品或同一样品测试值之间的变动性。

G.3.6 极小值、极大值、变动范围

极小值(x_{min}): 一组结果中的最小值或最低值。

极大值(x_{max}): 一组结果中的最大值或最高值。

变动范围(Range): 极大值与极小值之差, 即($x_{max} - x_{min}$) 误差分析中称之为极差。

G.3.7 标准四分位法双样品计算示例(excel)

标准四分位法双样品计算是一种可用 excel (或计算机编程) 实现的简单方法, 下面介绍一下如何使用 excel 公式函数及公式进行实验室间 z 比分数(z_b)和实验室内 z 比分数(z_w)计算及能力验证结果评定。

将某次能力验证中钢筋化学分析中的碳含量检测项目的各参加者试验结果录入到 excel 中, 如下表所示:

表 G2 四分位法双样品计算示例

	A	B	C	D	E	F	G
1	实验室代码	样品A	样品B	D	S	ZB	ZW
17	011	0.18	0.15	0.020506	0.235467	0.41	0.44
18	009	0.18	0.15	0.02192	0.22981	-0.19	0.68
19	007	0.18	0.15	0.021284	0.238224	0.70	0.57
20	016	0.19	0.17	0.015556	0.255973	2.59	-0.42
21	026	0.19	0.16	0.022627	0.241831	1.09	0.80
22	002	0.17	0.16	0.012021	0.232638	0.11	-1.04
23	020	0.18	0.14	0.027224	0.229739	-0.19	1.60
24	021	0.18	0.15	0.019799	0.234759	0.34	0.31
25	029	0.17	0.16	0.010607	0.232638	0.11	-1.28
26	037	0.20	0.15	0.03677	0.247487	1.69	3.26
27	028	0.18	0.17	0.012021	0.24678	1.61	-1.04
28	006	0.15	0.14	0.013435	0.204354	-2.89	-0.79
29	043	0.18	0.15	0.019092	0.235467	0.41	0.19
30	003	0.18	0.15	0.015556	0.231931	0.04	-0.42
31	018	0.17	0.15	0.014142	0.226274	-0.56	-0.67
32	034	0.17	0.14	0.017678	0.221324	-1.09	-0.06
33	031	0.17	0.15	0.016263	0.222739	-0.94	-0.30
34	041	0.17	0.15	0.018314	0.227618	-0.42	0.06
35	042	0.18	0.15	0.019799	0.231931	0.04	0.31
36	017	0.18	0.15	0.020506	0.231224	-0.04	0.44
37	035	0.18	0.16	0.019092	0.239709	0.86	0.19
38	032	0.18	0.16	0.012728	0.244659	1.39	-0.91
39	008	0.17	0.15	0.013435	0.22981	-0.19	-0.79
40	025	0.20	0.13	0.055084	0.234123	0.27	6.43
41	038	0.18	0.14	0.026163	0.221324	-1.09	1.42
42	010	0.18	0.14	0.025456	0.223446	-0.86	1.29
43	024	0.17	0.16	0.013435	0.232638	0.11	-0.79
44	040	0.17	0.14	0.017678	0.218496	-1.39	-0.06
45	036	0.17	0.15	0.015556	0.227688	-0.41	-0.42
46	结果数	44	44	44	44		
47	中位值	0.176	0.150	0.018	0.232		
48	nIQR	0.00908	0.00599	0.00577	0.00944		
49	稳健CV%	5.174	4.004	32.04	4.074		
50	最小值	0.151	0.127	0.006	0.204		
51	最大值	0.205	0.170	0.055	0.256		
52	极差	0.054	0.043	0.049	0.052		
53							

上表各单元格所用 excel 函数及公式如下：

单元格 B3:C45 录入结果值；

$E3=(B2+C2)/POWER(2,0.5)$;

$F3=(E2-SES47)/SES48$;

$D3=IF(MEDIAN(B2:B45)>MEDIAN(C2:C45),(B2-C2)/POWER(2,0.5),(C2-B2)/POWER(2,0.5))$;

$G3=(D2-SDS47)/SDS48$;

以上各列分别选中第一个数值单元格，下拉即可计算相应参数；

$B46=COUNT(B2:B45)$

B47=MEDIAN(B2:B45);

B48=(QUARTILE(B2:B45,3)-QUARTILE(B2:B45,1))*0.7413;

B49= B48/B47*100;

B50=MIN(B2:B45);

B51=MAX(B2:B45);

B52= B51-B50, 选中 B46:B52 后向右拖动即可计算各列统计量。

G.3.8 实验室间实验室间 z 分数 (zb) 和实验室内 z 比分数 (zw) 柱状图示例

用 Excel 中的图表功能作图如下:

- (1) 选中表 1 数据中的 A3:G45 单元格;
- (2) 点击数据下拉菜单中的排序, 出现“排序”对话框, 在“主要关键字”中选“zb”, 再选“升序”, 按确定;
- (3) 用 CTRL 键同时选中 A2:A45 和 F2:F45 单元格。点击插入, 选“柱形图”, 插入后调整坐标轴和柱状图图形参数, 在“标题”选项中的“图表标题”中填入“***的 zb 柱状图”。
- (4) zw 柱状图画法同上。

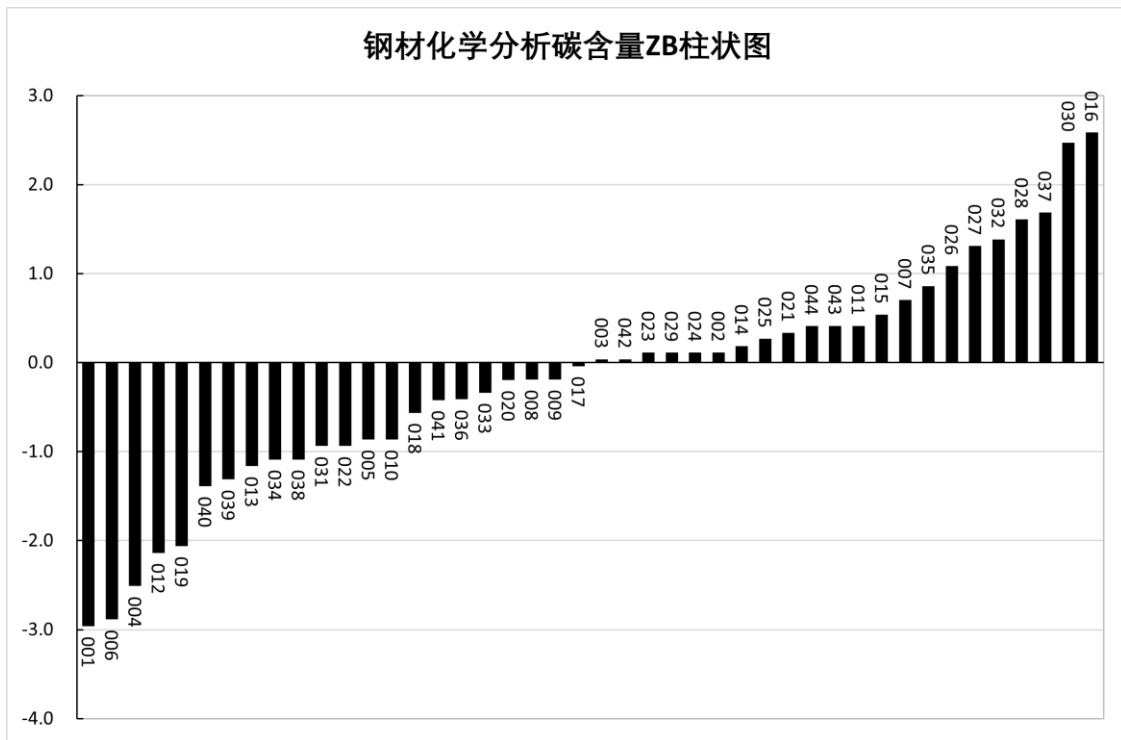


图 G2 zb柱状图

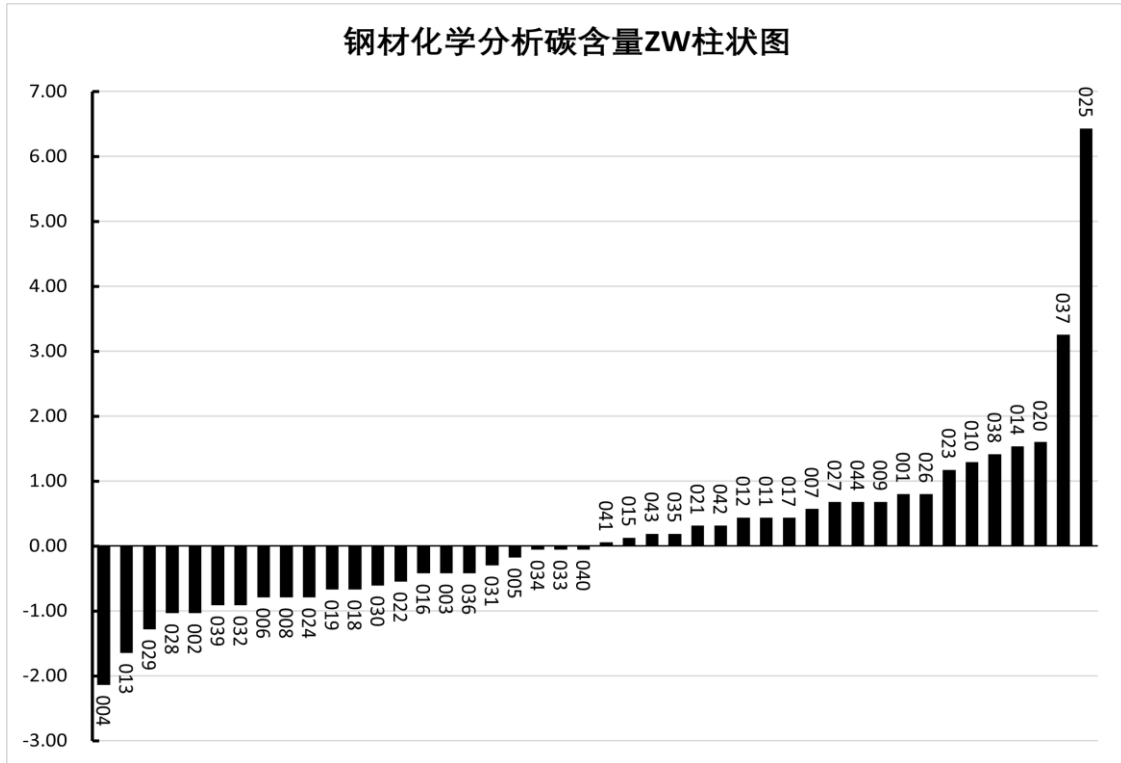


图 G3 zw柱状图

G.3.9 尧敦图示例

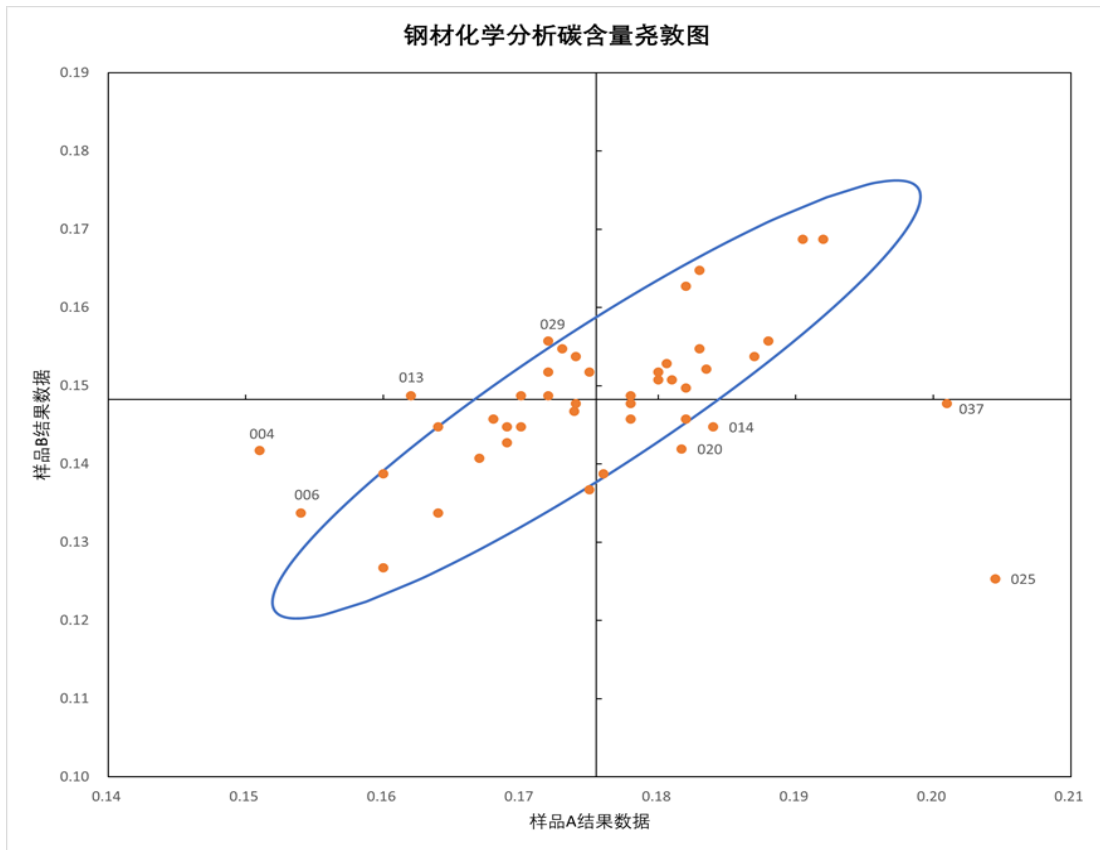


图 G4 尧敦图

G.4 参加者较少时的方法

G.4.1 如果能力验证参加者较少（少于 18 家）时，一般不能直接采用由参加者结果得到的公议值作为指定值对参加者进行评价。理想状态下宜使用独立于参加者的有效计量程序确定指定值，例如，根据配方或由参照实验室确定。能力评定准则也宜基于外部标准确定，例如专家判断，或基于目标适用性的标准。在这些理想情形下，使用预先确定的指定值和能力评定标准差来评定能力，因此，可执行只有一个参加者的能力验证。

G.4.2 当参加者较少且不能使用独立于参加者的、有明确溯源途径的方法确定指定值时，指定值或/和能力评定标准差可能仍需要利用参加者公议值得到。此时如果使用稳健统计方法，则应在考虑离群值的影响后，为能力评定设定参加者的最低数量。

- 1 当参加者数量 $p > 12$ 时可利用 G.3 算法 A 计算稳健平均值；
- 2 当参加者数量 $p = 2$ 时，可使用中位数（当其等于均值）作为稳健平均值；
- 3 其它情况可参考 GB/T28043-2019 附录 D。

G.4.3 当能力验证参加者数量 p 较少时，能力评定标准差的确定建议如下：

- 1 $p = 2$ 时：可使用 $\frac{|x_1 - x_2|}{\sqrt{2}}$ 确定稳健标准差；
- 2 GB/T28043-2019 附录 C.5.2 中标准差的 Q_n 估计；
- 3 根据与中位数的绝对距离得到稳健标准差：

$$s^* = \frac{1}{0.798 \times p} \sum_{i=1}^p |x_i - \text{med}(x)| \dots\dots\dots (G.22)$$

式中： s^* 是结果的稳健标准差；
 p 为参加者（实验室）数；
 x_i 为第 i 个参加者（实验室）结果数据；
 $\text{med}(x)$ 为中位数。

G.4.4 有关少量参加者进行能力验证的更多技术指南参见 GB/T28043-2019 附录 D 的 D.1。

G.5 混凝土结构实体强度现场能力验证统计方法

G.5.1 指定值的确定

1 由专家参加者（专家实验室）结果确定的公议值作为指定值。专家参加者应具有可证实的测定被测量的能力，并使用已确认的、有较高准确度的方法。

2 将参加者结果确定的公议值作为指定值。当采用稳健统计方法时，在考虑离群值的影响后，参加者数量不宜小于 18 个。

3 离群值的统计处理同 8.3.3。

注：混凝土结构实体强度检验能力验证的指定值通常应由满足要求的技术专家公议确定，技术专家的数量不少于 3 人。技术专家的选择宜遵循以下原则：

- (1) 具有高级及以上技术职称，且在行业内被普遍认同；
- (2) 从事混凝土结构实体强度相关检测或检验工作不少于 10 年，熟悉相关标准，技

术规范；

(3) 理解能力验证的目的、运作过程和结果利用情况，熟悉检验检测机构资质和能力评价相关要求。

G.5.2 能力评定标准差的确定

能力评定标准差可由以下方法确定：

- 1 与能力评价的目标和目的相符，由专家判定或法规规定（规定值）；
- 2 根据以前轮次的能力验证得到的估计值或由经验得到的预期值（经验值），应按照

GB/T28043-2019 中 C.4 的算法 S；

- 3 由参加者结果得到的稳健标准差、标准化四分位距等。其中，稳健标准差应按照 G.3.2 给出的算法 A 计算；标准化四分位距按式（G.16）计算。

G.5.3 有关混凝土结构实体强度现场能力验证统计方法的更多技术指南参见 RB/T145-2018。

本导则用词说明

1 为便于在执行本导则条文区别对待，对要求严格程度不同的用词说明如下：

1) 表示很严格，非这样做不可的用词：

正面词采用“必须”，反面词采用“严禁”；

2) 表示严格，在正常情况下均应这样做的用词：

正面词采用“应”，反面词采用“不应”或“不得”；

3) 表示允许稍有选择，在条件许可时首先这样做的词：

正面词采用“宜”，反面词采用“不宜”；

4) 表示有选择，在一定条件下可以这样做的，采用“可”。

2 本导则中指明应按其他有关标准、规范执行的，写法为“应符合……的规定”或“应按……执行”。

引用标准名录

- 1 《合格评定能力验证的通用要求》 GB/T 27043
- 2 《利用实验室间比对进行能力验证的统计方法》 GB/T 28043
- 3 《能力验证结果的统计处理和能力评价指南》 CNAS-GL 002
- 4 《能力验证计划的选择与核查及结果利用指南》 RB/T 031
- 5 《混凝土结构实体强度能力验证实施指南》 RB/T 145
- 6 《实验室能力验证 第2部分：名词和术语》 SN/T 2723.2
- 7 《测量不确定度评定与表示》 JJF 1059.1
- 8 《数值修约规则与极限数值的表示和判定》 GB/T 8170

山东省建设工程质量检测能力验证技术导则

条文说明

目 次

1	总则.....	61
2	术语.....	62
3	基本规定.....	64
4	能力验证的组织者.....	65
5	能力验证的承担者.....	66
6	能力验证的参加者.....	68
7	能力验证样品.....	69
8	能力验证结果统计与能力评价.....	70
9	能力验证结果处理.....	72
10	能力验证服务管理平台.....	73

1 总则

1.0.1 本条是本导则编写的宗旨和目的，以规范山东省建设工程质量检测能力验证活动的组织实施，促进能力验证工作的科学性、规范性、有效性，充分发挥能力验证作为评价我省建设工程质量检测机构能力的作用。同时为各级住房城乡建设主管部门开展能力验证工作提供了必要的参考。

1.0.2 本条界定了本导则的适用范围，适用于在山东省行政区域内的建设工程质量检测能力验证工作和后续监督管理工作。

1.0.3 建设工程质量检测能力验证除了执行本导则，还应遵守国家、行业及地方现行有关标准和规定。

2 术语

本章列出 19 个常用术语，以简化和规范本导则条文，使用更方便、精炼、表达意思更一致。这些术语是针对本导则定义的，其他地方使用仅供参考。

2.0.3 能力验证（也称水平测试），是通过实验室间检测结果的比对来判定实验室能力的合格评定活动。它通过发送统一制作的测试样品给各个实验室进行实际测试（测量），再将实验室的测试结果进行统计分析，通过各个实验室结果的一致性来判定实验室对于特定项目的检测能力。能力验证是对实验室能力状况和管理状况进行客观考核的一种方法，通过开展能力验证，可发现实验室存在的问题和监控实验室的运行状态，提高实验室检测能力和检测水平，确保检测的质量。能力验证一词的范围极为广泛，其实施模式包括诸如顺序参加模式、同时参加模式、分割样品测量模式、定性测量模式、已知值模式、部分过程模式和连续模式。能力验证工作有关的国家标准、行业标准有很多，如 GB/T 28043、GB/T 27043、SN/T 2723.2、RB/T 031、CNAS-GL002 等，国家及各地针对能力验证也出了一些管理办法和行政性文件，能力验证的组织者应根据这些标准及规定制定相应的能力验证管理要求。

2.0.4 能力验证的组织者指的是山东省、16 个设区市、县级住房城乡建设主管部门。省住房城乡建设主管部门负责统一协调、组织实施、监督管理全省检测能力验证工作。设区市、县级住房城乡建设主管部门负责本行政区域内检测能力验证的统一协调、组织实施和监督管理工作。

2.0.5 除能力验证的承担者外，很多标准中还包括能力验证的提供者（对能力验证计划建立和运用中的所有任务承担责任的组织），本导则中的能力验证的承担者包括能力验证的提供者。

2.0.8 稳健统计方法是近年来数理统计中倍受重视的一个重要分支，它对极端结果的处理不是将其从数据组中剔除，而是给予其较小的权，使其对平均值估计值和标准差估计值的影响减至最小，其稳健性主要是指统计方法对于偏离假定模型的不敏感性。事实上，任何随机现象都难以用一个简单的数学模型准确地描述。当实际观测数据中包含异常点而明显不对称，或当总体的实际分布不对称时，应用传统的统计方法有可能会得出错误的结论，而采用稳健统计方法则可以有效地克服观测数据偏离独立、正态分布等假设或包含异常点时对结果判断的困难。

2.0.11 公议值通常宜用于描述一轮能力验证中参加者结果的均值和方差，但也可以仅用于部分参加者结果，可以是参加者结果，也可以是专家实验室结果，比如，仅根据若干专家实验室结果得到的公议值。

2.0.12 离群值可能来源于不同的总体，或由于不正确的记录或其他粗大误差的结果。

2.0.13 标准差只适用于比例尺度和定距尺度的结果。并非所有的能力验证都根据结果的分散性进行评价。

2.0.15 z 比分数有时也称为 z 值或 z 分数。当采用不同形式的指定值和能力评价标准偏差或统计方法时，在 z 比分数前冠以相应的修饰词。如实验室内 z 比分数、实验室间 z 比分数、稳健 z 比分数等。能力验证中可能导致结果不一致的来源主要有检测样品的差异、实验室间的差异（包括检测方法间的差异）和实验室内部的差异。但由于能力验证使用的检测样品一般已通过规定的程序预先确认其量值充分均匀，因此，实验室结果的差异主要来自实验室间和实验室内这两方面的差异。为了评审实验室间和实验室内这两种差异性，参加实验室必须进行多于一次（如两次）的相同检测，分割水平检测样品也正是因此提出来的。分割水平样品，是指用于能力验证的样品为两个量值相近、但不相同的样品，检测机构可以采用相同的检测条件测定其量值。在能力验证中采用分割水平检测样品，既可以客观地评价参加实验室结果与指定值的偏离情况，即实验室系统误差，也可以客观地评价参加实验室在被测量水平下数据分散情况，即实验室随机误差。同时，分割水平检测样品也避免了用同一检测样品作重复测量，或在同一轮能力验证中，包含两个完全一致样品所伴随的问题。分割水平检测样品已成为当前最常用的实验室间检测能力验证方法。对于分割水平样品的检测结果来说，稳健 z 比分数分为实验室间 z 比分数（ zb ）和实验室内 z 比分数（ zw ）。对于单一检测样品的能力验证来说，一个简单的稳健 z 比分数的符号表明结果太高（正 z 比分数）或太低（负 z 比分数），但不能区分实验室的两种差异，也不能得知结果异常是由于实验室间系统误差还是由于实验室内随机误差，或者是两者共同造成的。对于分割水平检测样品的能力验证来说， zb 的大小代表某实验室的 S 与中位值 M （ S ）的偏离程度，反映了实验室测试结果的系统误差，而符号“+”、“-”代表与 M （ S ）的偏离方向，由于采用加和的形式，因此对结果偏离的反应较单个样品灵敏； zw 的大小代表某实验室的 D 与中位值 M （ D ）的偏离程度，反映了实验室测试结果的随机误差，或者说是实验室测试系统的稳定性。正的实验室间离群值（ $zb > 3$ ），表明结果都太高；负的实验室间离群值（ $zb < -3$ ），表明结果都太低。正的实验室内离群值（ $zw > 3$ ），表明结果间的差值太大，负的实验室内离群值（ $zw < -3$ ），表明结果间的差值太小或相对于中位值是在“相反的方向”。

3 基本规定

3.0.1 本条规定了能力验证的原则，目的是为了保证能力验证的实施效果，同时避免不必要的资源浪费，是保证能力验证工作质量的原则。

3.0.2 被委托的机构应具备本导则第 5.1.2 条规定的的能力。能力验证的组织者在开展能力验证活动前应依据第 5.1.2 条内容对能力进行确认，拟委托机构达不到要求不应委托其承担能力验证工作。

3.0.5 本条中规定的的能力验证的参加者是指由能力验证的组织者根据已取得资质证书的检测机构名单确定的、具有相应检测能力的检测机构，已确定的具有相应资质的检测机构必须参加能力验证，本条规定的参加者不包括未取得相应检测资质证书自愿参加能力验证的检测机构。

3.0.7 本条指的能力验证的组织者根据其管理需要，可组织或不再组织第二次能力验证，如组织开展第二次能力验证，则程序与第一次能力验证完全相同。

4 能力验证的组织者

4.0.1 本导则中的能力验证组织者指的是省、设区市、县三级住房城乡建设主管部门，各级住房城乡建设主管部门负责职责范围和辖区内的能力验证的统一协调、组织实施、监督管理工作。

4.0.3 本条所指的涉及结构安全的项目包括（但不限于）水泥、混凝土、砂浆、外加剂、钢筋、混凝土实体结构检测及地基基础检测等。

4.0.4 与其他行业组织能力验证不同，我省住房城乡建设行业在确定能力验证参加者时不需要参加者自行报名，而是根据能力验证的项目和行政区划内取得相应资质证书的检测机构名单来确定，如某市住建局在组织混凝土实体结构检测能力验证时，只需要确定该市取得主体结构检测资质证书的检测机构名单即可确定参加者名单，然后通知其参加相应的能力验证活动。

5 能力验证的承担者

5.1 一般规定

5.1.1 本条规定了能力验证的承担者所负责的具体工作，同时能力验证活动应符合相关技术标准的规定。

5.1.2 本条规定了能力验证的承担者应满足的基本条件，能力验证的组织者在选择确定能力验证的承担者时应符合本条要求。

5.1.7 能力验证的承担者应提高自身工作能力，保证工作质量，采取相应措施规避消除来自内外部的、不正当的商业、财务和其他方面的压力和影响，确保能力验证结果准确、客观、准确、公正。

5.2 人 员

5.2.2 建议考虑人员定期再培训需求。人员培训政策宜考虑技术更新和证明持续能力的需求，并以持续提高技能为目的。

5.3 设备、设施和环境

5.3.4 本条中的环境条件包括诸如消毒、灰尘、电磁干扰、辐射、湿度、供电、温度、声级、振级，要与涉及的技术活动相适宜。

5.4 能力验证实施方案

5.4.1 能力验证承担者在建立一项计划及相关信息时可考虑相关方利益。

5.4.2 能力验证承担者应获得在检测、校准、抽样、检查以及统计分析等相关领域的必要专业知识和经验。需要时，可通过成立顾问组（称谓可不同）来实现。

5.5 能力验证的运作

5.5.2 6 作业指导书通常包括计量单位、有效数字或小数位数、报告结果的基准等参数。

5.6 技术报告

5.6.1 在技术报告中，如果不能为参加者提供所有原始资料，那么可以提供一个结果汇总，例如，以表格或图形的形式。

5.6.2 对于连续能力验证，提供较简单的报告即可，本条中很多内容在常规报告中可以省略，但应包含在参加者可获得的能力验证协议或阶段性的汇总报告中。

6 能力验证的参加者

6.0.1 参加能力验证的检测机构（参加者）应当正确认识能力验证活动目的和意义，按照住房城乡建设主管部门要求参加能力验证活动，通过能力验证确定自身的检测能力。

6.0.2 能力验证的参加者在得到参加能力验证的通知后，在收到能力验证样品前，应提前做好各项准备工作，确认人员、设备标准物质、环境条件等是否满足要求，同时可开展模拟试验，以发现各环节存在的问题并及时解决，为能力验证试验打好基础。

6.0.3 能力验证的参加者应认真细致地阅读作业指导书，不能有遗漏事项，以免由于未按说明盲目进行操作或计算导致结果的错误或偏差。能力验证的参加者若私下比对串通能力验证数据、结果或出具虚假能力验证数据、结果，在某种程度上会对能力验证的结果产生一定的影响。能力验证的承担者可通过设计分割水平样品、减少样品发放量、缩短能力验证结果上报周期、提供检测影像资料等具体措施，有效地防止能力验证数据的串通，保证能力验证活动的健康发展。

6.0.4 对能力验证结果存在异议的，参加者可以在得知能力验证结果之日后及时向能力验证的组织者提出申诉。能力验证的组织者应组织专家和承担机构对申诉内容进行研究，并及时给出答复。除申诉外，参加者对能力验证的意见建议也应及时提出，便于能力验证的组织者改进和完善能力验证工作。

6.0.5 本条规定结果可疑或离群的，要从“人、机、料、法、环、溯、测”等方面查找原因，采取有效措施积极整改，并在采取有效纠正措施后报名参加第二次能力验证。造成结果不满意或可疑的原因，主要有管理和技术上的原因，管理原因如：抄写错误、贴错标识、小数点错误等。技术上的原因如：样品的储存或处理不当、检测方法或内部质量控制有问题、标准物质/标准样品异常、设备状态不佳、环境条件不适宜或数据处理出现问题等。通常采取的纠正措施有：

- 1 核查相关人员是否理解并遵循检测程序；
- 2 核查检测程序的所有细节是否正确；
- 3 核查设备校准和试剂的成分；
- 4 更换可疑的设备或试剂；
- 5 与另一个检测机构进行人员、设备和/或试剂的比对试验。

7 能力验证样品

7.0.1 对能力验证的组织者来说，如何制造或挑选出满足能力验证需求的样品，是能力验证成败的关键。如何在一组样品测试数据中，利用数理统计的理论进行筛选，确保样品均匀性；在检测过程中，尽可能减少样品受系统误差的影响，导致能力验证结果的偏离，为参与活动的检测机构测试能力提供依据。能力验证的承担者应考虑所有与被测样品有关的、可能影响能力验证完好性的条件，如均匀性、抽样、稳定性、在传运过程中可能的损坏及周围环境条件的影响等，避免样品受到不利的影响。

7.0.5 能力验证的承担者可选择具有资质的服务方将能力验证样品分发至参加者，并确认送达情况。

7.0.6 为确保参加者能及时接收到完整的样品以及减少在样品接收状态方面的纠纷，能力验证的参加者应按照作业指导书的要求及时对样品状态进行确认，如发现样品有损坏或其他异常情况，应及时与能力验证的承担者联系。当出现参加者未收到样品或收到后发现样品存在缺陷不适合检测的，如样品损坏或丢失，应及时与承担者联系，实施补救措施。

8 能力验证结果统计与能力评价

8.2 统计方法的选择

本条规定了统计方法的选择。能力验证中使用的统计方法需要适合能力验证的目的，符合统计原理。在能力验证的实际操作中，统计方法的选择要综合运用统计知识以及专业的判断，以获得参加者的准确评价。

除了四分位法等传统稳健统计方法外学界还进行了一些统计学研究方法，有主成分分析方法、非负矩阵分解方法、核密度估计方法和 Bootstrap 方法等。但应用最多的稳健统计方法是基于能力验证结果的中位数和四分位距，它们是数据中心和离散性的度量，类似于经典统计方法的均值和标准差。

对于结果数据的对称、单峰分布，四分位法和算法 A 可以较好估计均值和标准差，但对于大比例 (>20%) 离群值，或者有小众数的情形 Q/Hampel 方法则更为适用，可参考 GB/T28043-2019 附录 C.5。

如果能力评定准则采用过去或本轮能力验证的公议值，则能力评定标准差可使用参加者结果标准差的稳健估计，此时使用 z 比分数等能力评定统计量和能力评定标准差进行能力评定最为简便。用参加者结果确定标准差主要优点在于简单和实用，可用于很多场合，且可能是唯一可行的方法。需要注意的是，不同的稳健统计方法具有不同的适用条件、统计效率和失效点。

8.4 能力验证结果的统计处理

8.4.3 能力评定的解释基于正态分布假定，但只需合格参加者的结果满足正态分布即可。通常情况下没有必要确认测量结果是正态分布的，但有必要确认其近似对称分布，至少在视觉上呈现为对称的。一般宜通过数据直方图或核密度图核查结果分布是否对称，统计假设是否合理，是否存在异常（如双峰分布、离群值比例较大或异常偏倚）等。

核密度图以一条光滑的曲线来描述数据分布的一般形状，曲线形状表明了数据的分布情况，多个不同的众数会形成分离的峰；边缘值也会形成与数据主体显著分离的峰。因此对于结果分布为不对称分布，先用核密度图识别多峰性，再选择适用于不对称结果的稳健技术。

直方图或核密度图最好使用中（参加者大于 50 小于 100）到大规模（参加者大于 100）数据集，因为小数据集（10 个或更少）有可能包含轻微离群的值或明显的众数，尤其是基于稳健标准差计算带宽时。

8.4.6 当参加者数据中有离群值时，宜优先采用稳健统计，但对于数据量较少的情况通常并不推荐使用。对于非常小（如参加者数量小于 18）的数据集，可能需要优先使用经典统

计方法，使用离群值检验，先剔除离群值，然后计算均值和标准偏差，而不同的离群值检验适用于不同规模的数据集。另外如果参加者的数量对于所使用的具体确定指定值或能力评定标准差的方法而言太少，能力评定可能会变得不可靠；因此，应考虑是否为能力评定设定参加者的最低数量。

8.6 能力评定

8.6.2 在稳健统计分析中常将各种比对情况用诸如柱状图、Youden（尧敦）图等图形展示出来。用图形来比对和分析各参加者的检测能力简明直观、易于理解和便于比较。为此，表格和图形在能力验证的稳健统计技术中应用十分广泛。最常用的图形是 z 比分数（ z, z_b 和 z_w ）序列柱状图和 Youden（尧敦）图。

9 能力验证结果处理

9.0.2 在能力验证的组织者确定了参加者的名单以后，参加者除非注销了或短期内计划注销相应资质等特殊情况以外，均应根据组织者的要求参加能力验证活动。

9.0.5 本条是对能力验证的承担者做出的约束，一旦承担者偏离了公正性，能力验证就失去了意义，社会影响会极为恶劣。

9.0.6 本条规定了能力验证活动的监督方式，能力验证活动的监督不仅仅是对参加者的监督，也是对承担者的监督，同时也能发现一些不良行为，保证能力验证活动能够客观公正地进行。

10 能力验证服务管理平台

10.0.1 目前很多地方能力验证的实施者是通过人工电话通知、邮寄样品、传真样品状态等传统人工操作的方式，工作量大且易出错，统计和查询不方便。能力验证服务管理平台充分利用互联网+技术，去除传统能力验证人工实施中的诸多弊端，能极大地方便能力验证参加者、组织者、承担者，缩短在时间上和精力上的投入，有效地提高结果的准确度和判定的效率，提升工作效率，优化服务效能。

10.0.6 本条规定平台应能对各项操作事项进行记录，目的是为了保证所有相关信息可追溯。