



# 中华人民共和国国家标准

GB/Z XXXXX—XXXX

## 化学分析实验室质量保证及质量控制指南

点击此处添加标准名称的英文译名

(点击此处添加与国际标准一致性程度的标识)

征求意见稿

XXXX—XX—XX 发布

XXXX—XX—XX 实施

国家市场监督管理总局  
国家标准化管理委员会 发布

## 目 次

1	范围.....	1
2	规范性引用文件.....	1
3	术语和定义.....	1
3.1	校准 calibration.....	1
3.2	检定 verification.....	1
3.3	计量标准（又称测量标准）measurement standard.....	1
3.4	核查标准 check standard.....	2
3.5	期间核查 intermediate checks.....	2
3.6	校准曲线 calibration curve.....	2
3.7	标准物质/标准样品 reference material (RM) .....	2
3.8	重复性测量条件 Repeatability Condition of Measurement.....	2
3.9	再现性测量条件 Reproducibility Condition of Measurement.....	2
3.10	能力验证 proficiency testing.....	2
3.11	测量审核 measurement audit.....	2
4	总则.....	2
5	监控结果有效性方案策划.....	3
5.1	监控目标.....	3
5.2	监控项目.....	3
5.3	监控方法.....	4
5.4	监控结果评价.....	4
6	内部控制方法.....	4
6.1	标准物质/标准样品或质控样品.....	4
6.2	校准曲线.....	6
6.3	加标回收.....	8
6.4	空白样品.....	9
6.5	控制图.....	9
6.6	设备期间核查.....	13
6.7	留样复测.....	15
6.8	重复分析.....	16
6.9	盲样测试.....	17
6.10	仪器比对.....	17
6.11	方法比对.....	18
6.12	人员比对.....	19
6.13	结果相关性.....	20
7	外部控制方法.....	21
7.1	能力验证.....	21
7.2	测量审核.....	21
7.3	实验室间比对.....	22
	附录 A（资料性） 数据处理.....	25

A.1 数据的统计处理.....	25
A.2.1.1.1 格拉布斯 (Grubbs) 检验.....	25
附录 B (资料性) 用 EXCEL 制作控制图.....	41
附录 C (资料性) 结果有效性监控活动策划思维导图示例.....	45
附录 D (资料性) 监控策划示意表.....	46

# 化学实验室 监控结果有效性指南

## 1 范围

本文件给出了化学分析实验室监控结果有效性的总则、方案策划、内部控制方法、外部控制方法和数据处理等信息。

本文件指导化学分析实验室对结果有效性的监控。

## 2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 8170-2008 数值修约规则与极限数值的表示和判定

GB/T 6379.2-2004 测量方法与结果的准确度（正确度与精密度）第2部分：确定标准测量方法复性与再现性的基本方法

GB/T 6379.5-2006 测量方法与结果的准确度（正确度与精密度）第5部分：确定标准测量方法密度的可替代方法

GB/T 27407-2010 实验室质量监控 利用统计质量保证和控制图技术 评价分析测量系统的性能

GB/T 32464-2015 化学分析实验室内部质量监控 利用控制图核查分析系统

GB/T 32465-2015 化学分析方法验证确认和内部质量监控要求

GB/T 32467-2015 化学分析方法验证确认和内部质量监控 术语及定义

ISO 11843-1:2008 检测能力 第1部分：术语和定义

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1 校准 calibration

在规定条件下的一组操作，其第一步是确定由测量标准提供的量值与相应示值之间的关系，第二步是用此信息确定由示值获得测量结果的关系，这里测量标准提供的量值与相应示值都具有测量不确定度。

【JJF 1001-2011，定义 4.10】

### 3.2 检定 verification

查明和确认计量器具是否符合法定要求的活动，它包括检查、加标记和/或出具检定证书。

【JJF1001-2011，定义 9.17】

### 3.3 计量标准（又称测量标准） measurement standard

具有确定的量值和相关联的测量不确定度，实现给定量定义的参照对象。

【JJF 1033-2016，定义 3.1】

### 3.4 核查标准 check standard

用于日常验证测量仪器或测量系统性能的参照对象。

### 3.5 期间核查 intermediate checks

根据规定程序,为了确定计量标准、标准物质或者其他测量仪器是否保持其原有状态而进行的操作。

【JJF 1001-2011, 定义 9.49】

### 3.6 校准曲线 calibration curve

表示示值与对应测得值关系的曲线。

注:校准曲线表示了一对一的关系,由于它没有关于测量不确定度的信息,因而没有提供测量结果。

### 3.7 标准物质/标准样品 reference material (RM)

具有一种或多种规定特性足够均匀且稳定的材料,已被确定其符合测量过程的预期用途。

注1:标准物质/标准样品是一个通用术语。

注2:特性可以是定量的或定性的(例如:物质或物种的特征属性)。

注3:用途可包括测量系统的校准、测量程序的评估、给其他材料赋值和质量监控。

注4:ISO/IEC Guide 99:2007(5.13)有类似的定义,但限定“测量”术语仅用于定量的值。然而,ISO/IEC Guide 99:2007(5.13)的注3中明确包括定性特性,称作“标称特性”。

### 3.8 重复性测量条件 Repeatability Condition of Measurement

在同一实验室,由同一操作者使用相同设备,按相同的测试方法(测量程序),在短时间内对同一被测对象相互独立进行的测试条件。

### 3.9 再现性测量条件 Reproducibility Condition of Measurement

在不同的实验室,由不同的操作者使用不同的设备,按相同的测试方法(测量程序),从同一被测对象相互独立进行的测试条件。

注:对再现性的表述应指出相关测试条件的变化。再现性可用测试结果的分散性量化表述。

### 3.10 能力验证 proficiency testing

能力验证利用实验室间比对,按照预先制定的准则评价参加者的能力。

【GB/T2704-2019 3.2】

### 3.11 测量审核 measurement audit

一个参加者对被测物品(材料或制品)进行实际测试,其测试结果与参考值进行比较的活动。注:测量审核是对一个参加者进行“一对一”能力评价的能力验证计划

## 4 总则

4.1 实验室宜制定针对检测结果有效性监控的文件。

4.2 实验室对结果的监控宜覆盖到所有检测项目,确保检测结果的准确性和稳定性。

- 4.3 实验室定期对监控进行策划，形成年度质控方案。
- 4.4 实验室分析监控活动的数据用于控制实验室活动，适用实施改进。
- 4.5 若发现监控活动数据分析结果超出预定的准则时，应采取适当措施防止报告不正确的结果。
- 4.6 实验室制定内部质量监控方案时应考虑以下因素：
- 检测或校准业务量；
  - 检测或校准结果的用途；
  - 检测或校准方法本身的稳定性与复杂性；
  - 对技术人员经验的依赖程度；
  - 参加外部比对（包含能力验证）的频次与结果；
  - 人员的能力和经历、人员数量及变动情况；
  - 新采用的方法或变更的方法等。
- 4.7 实验室可采取多种适用的质量监控手段，但不限于：
- 定期使用标准物质、核查标准或工作标准来监控结果的准确性；
  - 通过使用质量监控图来监控结果的准确性和精密度；
  - 通过获得足够的标准物质，评估在不同浓度下检测结果的准确性；
  - 定期留样再测或重复测量以及实验室内比对，监控同一操作人员的精密度或不同操作人员间的精密度；
  - 采用不同的检测方法或设备测试同一样品，监控方法之间的一致性；
  - 通过分析一个物品不同特性结果的相关性，以识别错误；
  - 进行盲样测试，监控实验室日常检测的准确度或精密度水平。
- 4.8 一些特殊的检测活动，检测结果无法复现，实验室应关注人员的能力、培训、监督以及与同行的技术交流。
- 4.9 外部质量监控方案优先考虑适合的能力验证计划和测量审核，适当时，可组织实验室间比对计划。
- 4.10 实验室制定外部质量监控计划应考虑以下因素：
- 内部质量监控结果；
  - 实验室间比对（包含能力验证）的可获得性，对没有能力验证的领域，实验室应有其他措施来确保结果的准确性和可靠性；
  - 客户和管理机构对实验室间比对（包含能力验证）的要求。

## 5 监控结果有效性方案策划

### 5.1 监控目标

确保结果有效性的监控（以下简称“结果监控”）是实验室为持续监控测量过程和测量结果以确定结果是否足够可靠，能达到可以发布的程度而采取的一组操作。实验室应持续稳定地将监控活动纳入技术运作和管理体系运行的重点工作中，并制定监控目标。监控目标应与标准要求、客户要求、实验室内部文件或特定项目的需求保持一致，以保证结果的有效性。

### 5.2 监控项目

策划的监控计划应覆盖到实验室的所有检测项目，并能有效监控检测结果的准确性和稳定性。制定内部结果监控计划应考虑检测业务量、检测结果的用途，并重点关注以下检测项目：

- 新开展或方法变更的项目；
- 重要的、技术难度高的检测项目；
- 前期内部及外部质控结果有问题及不满意的检测项目；

- 对技术人员经验的依赖程度强的检测项目；
- 检测频次少的项目（如上个质控周期少于1次）

### 5.3 监控方法

5.3.1 为保证确保结果有效性的监控的有序进行，实验室结果监控活动需要进行长期整体的策划，一般以监控计划的形式输出可通过内部结果监控和外部结果监控相结合的方式进行，相关示例见附录D。

5.3.2 内部质量监控方法包括质控样品、校准曲线核查、加标回收、空白样品、控制图、设备期间核查、留样复测、重复分析、盲样测试、仪器比对、方法比对、人员比对、结果相关性等。

5.3.3 外部质量监控方法包括能力验证、测量审核、实验室间比对等外部质量监控方法以及外部机构组织的其他质量监控活动。

5.3.4 附录C给出了结果有效性监控活动策划思维导图示例，附录D假定了实验室检测能力后输出的策划计划。由于化学涉及行业的多样性，检测特性的复杂性，示例只给实验室提供一个可供参考的策划思路，不能作为模板套用。

### 5.4 监控结果评价

对监控结果进行评价时，其判定标准应满足相应检测方法中的质量保证与质量监控规定，方法未作要求或要求不充分时，可以由实验室结合相关检测方法要求、仪器设备检定/校准证书、标准物质证书作为质控结果的判定标准，或者根据统计方法进行判断。附录D给出了监控策划示意图。

## 6 内部控制方法

### 6.1 标准物质/标准样品或质控样品

#### 6.1.1 标准物质/标准样品或质量监控物质的用途

6.1.1.1 标准物质/标准样品（RM）或质量监控物质（QCM）可用于测量过程的各个阶段，主要用于评估测量系统和测量过程的精密度和稳定性。例如使用RM可以（但不限于）：

- 方法确认/验证中精密度评估；
- 测量过程中的质量监控；
- 运用相同检测方法测量相同RM，可将多次测量结果用于质量监控图的绘制（见6.5）；
- 实验室内和实验室间比对核查样品；
- 核查仪器稳定性；
- 评估环境条件（如温度，湿度）影响变化对测量结果的影响。

6.1.1.2 有证标准物质/有证标准样品（CRM）除适用5.1.1.1所列用途外，主要用于：

- 校准；
- 建立计量溯源性；
- 方法确认/验证中的正确度评估；
- 为其它材料赋值。

#### 6.1.2 标准物质/标准样品或质量监控物质的选择和使用

##### 6.1.2.1 标准物质/标准样品的选择

- 选用适宜的CRM、RM，一般应考虑以下因素（适用时，但不限于）：
- RM的含量水平（浓度）与预期应用的水平相适应；

- RM 的基体与待测试样的基体应尽可能接近；
- 应以与待测试样相同的形态使用；
- 最小取样量满足检测方法要求；
- 标准物质的数量应满足整个实验计划的使用；
- 标准物质的使用应在其注明的有效期限之内，并符合贮存条件；
- 可接受的测量不确定度。

### 6.1.2.2 质控样品的选择

实验室选择（或制备）质量监控物质（QCM），一般为以下情形：

- 需要一个尽可能接近日常样品的标准物质/标准样品用于质量监控；
- 需要一个适用于日常的标准物质/标准样品来补充有证标准物质；
- 没有合适的标准物质/标准样品；
- 无需使用有证标准物质/标准样品的所有特性（即溯源性和规定特性值的不确定度）的场合；
- 出于成本的考虑。

### 6.1.2.3 标准物质/标准样品或质控样品的使用方法

6.1.2.3.1 使用 RM 进行质量监控，可以定期或不定期安排 RM 的检测，更为有效的方法是在每个样品分析批次中，按一定的数量，例如 5% 随机或均匀插入 RM，与检测样品同时检测，依据检测结果与标准值的比较，可以评估实验室的检测过程是否可接受。如果结论为不可控，实验室应采取相关行动。实验室通过多次重复检测一个 RM 的结果建立质控图，与预先建立起来的标准相比较，监控检测过程是否受控。

6.1.2.3.2 在方法确认和验证过程中，可采用该方法检测 CRM，依据检测结果与标准值的比较，建立该方法的正确度参数，评估是否满足预期目标和该方法规定的要求。在方法确认和验证过程中，可采用该方法检测 RM 和或 CRM，依据检测结果的变化范围或结果的标准偏差，建立该方法的精密度参数，评估是否满足预期目标和该方法规定的要求。

6.1.2.3.3 实验室开展仪器比对、人员比对、方法比对等各类质控活动，通过检测 RM，评估比对结果。

### 6.1.3 标准物质/标准样品或质量监控物质检测结果评价

#### 6.1.3.1 评价依据

实验室应依据相关标准，结合对检测数据的质量要求，制定标准物质/标准样品（RM）或质量监控物质（QCM）检测结果的可接受限，做为评价依据。

#### 6.1.3.2 正确度评估

正确度评估一般可采用如下评价方法（不限于）。

a) 实验室检测结果与 CRM 证书给出的特性量值之差应小于与该值相关联的不确定度，即公式(1)

$$|x_{meas} - x_{CRM}| \leq k \sqrt{u_{meas}^2 + u_{CRM}^2} \quad (1)$$

式中：

$x_{meas}$  为 CRM 的测量结果；

$x_{crm}$  为 CRM 证书中给出特性量值；

$u_{meas}$  为与 CRM 的检测结果相关的标准不确定度；



$U_{crm}$  为 CRM 特性量值的标准不确定度。

$k$  为包含因子，通常包含概率为 95% 时，取  $k = 2$ 。

- b) 如果使用的标准方法已给出重复性限 ( $r$ ) 和再现性限 ( $R$ ) 时，可依据临界值 ( $CD'$ ) 评价 CRM 检测结果是否可接受。临界值的 ( $CD'$ ) 计算见公式 (2) 和公式 (3)。小于等于  $CD'$  时，为满意结果 (可接受结果)，反之大于  $CD'$  时为不满意结果 (不可接受结果)。

$$|X_{meas} - X_{crm}| \leq CD' \quad (2)$$

式中：

$X_{meas}$  为 CRM 的检测结果；

$X_{crm}$  为 CRM 证书中给出特性量值；

$CD'$  为临界值 (可接受限)。

$CD$  按公式 (3) 计算：

$$CD = \frac{1}{\sqrt{2}} \sqrt{R^2 - r^2 \frac{(n-1)}{n}} \quad (3)$$

式中：

$R$  标准方法给出的再现性限；

$r$  标准方法给出的重复性限；

$n$  测定次数；

$CD$  一定概率下的临界值。

- c) 当相关标准给出允许限时， $CD'$  按公式 (4) 计算：

$$CD' = \sqrt{(\Delta/\sqrt{2})^2 + U_{CRM}^2} \quad (4)$$

式中： $\Delta$  为相应测试标准规定的允许差。

- d) 采用  $En$  值评价结果， $En$  值小于 1 为可接受值，反之  $En$  值大于 1 为不可接受值， $En$  值按公式 (5) 计算：

$$E_n = \frac{x - x_{CRM}}{\sqrt{U^2 + U_{CRM}^2}} \dots\dots\dots (5)$$

式中：

$X_{meas}$  为 CRM 的检测结果；

$X_{crm}$  为 CRM 证书中给出特性量值；

$U_{meas}$  为与 CRM 的检测结果相关的标准不确定度；

$U_{crm}$  为 CRM 特性量值的标准不确定度。

- e) 依据相关标准规定评价。

### 6.1.3.3 精密度评估

精密度评估见本文件 5.8:

## 6.2 校准曲线

## 6.2.1 校准曲线的制作

校准曲线是检测过程中检测数据质量监控的重要参数，基体效应对分析结果有影响，通过在空白样品中加入已知浓度待测物制备校准曲线系列浓度的校准标样。

- a) 校准曲线的浓度范围由最低浓度和最高浓度来决定。最低浓度需要当接近检测方法报告限的水平（最低定量限），最高浓度不宜引起残留效应。需要涵盖相应的待测样品的浓度范围，并需进行与待测样品完全相同的分析处理。
- b) 校准曲线的不同浓度校准标样设置数需要符合方法标准或者专业技术领域的技术指导原则要求，或者结合仪器性能、待测物浓度及精密度要求等因素来具体确定。一般不少于5个系列浓度点。
- c) 校准曲线的系列浓度校准标样的准确度和精密度需符合方法标准或者专业技术领域的技术指导原则要求。
- d) 校准曲线系列浓度的校准标样的分析要合理放置在待测样本的分析批中，一般一个分析批开始和结束各分析一组校准曲线的不同浓度校准标样，或者至少一个校准曲线的中间浓度校准标样。
- e) 空白样品原则上需要与待测样品基体一致或者相似，如果不存在干扰或基质效应，也可以使用单一来源的空白样品。
- f) 校准曲线制作使用的标准物质需要确保满足分析方法的要求，首选有证标准物质，必要时要进行核查。当使用无证标准物质时，需要做必要的验证或者比对验证。
- g) 校准曲线的系列浓度校正样品的制备需采用不同的储备液或者储备物质制备，并且不能用于制备质控样品，以避免出现与方法分析性能无关的偏差。如果储备液或者固体标准物质的准确度和稳定性已得到验证，则可使用相同的储备液或者储备物质制备校准标样。
- h) 为保证准确测量，待测样品的浓度需要合理分布在校准曲线的浓度范围内，避免集中在校准曲线的一端，或者超出校准曲线范围，如出现上述情况，需要重新制作校准曲线，或者采取其他方式处理。

## 6.2.2 校准曲线的评价

### 6.2.2.1 综述

校准曲线直接影响样品分析结果，一般校准曲线为 $y = a + bx$ ，对校准曲线的相关系数 $r$ ，截距 $a$ 和斜率 $b$ 三个参数进行检验，以及对校准曲线的核查，可以评价分析结果的质量，保证分析结果准确性。

### 6.2.2.2 相关系数 $r$ 检验

相关系数（ $r$ ）表示校准曲线的精密度，相关系数的具体要求与待测浓度高低和检测方法等因素相关，例如分光光度法要求  $r \geq 0.999$ ，而一些基质复杂样本分析，相关系数要求被适当放宽，例如血浆中药物浓度测定， $r \geq 0.99$ 即可。如果 $r$ 不满足要求，校准曲线不能用于定量分析，应查找原因加以纠正，重新制作合格的校准曲线。

### 6.2.2.3 截距检验

#### 6.2.2.3.1 $t$ 检验

分析误差越小，截距越趋近于0。理想的校准曲线，截距 $a=0$ 。但由于存在难以控制的随机因素，分析中存在系统误差，多数校准曲线的截距不为0。当 $a \neq 0$ 时，将截距与0作 $t$ 检验，即显著性检验。经检验，在给置信水平下（通常取95%）无显著性差异时，可作0处理，表明标准曲线的回归计算结果准确

度满足要求。若截距与0存在显著性差异，表示校准曲线的回归计算结果准确度不高，不能进行结果回归计算，应找出原因予以纠正。

#### 6.2.2.3.2 剩余标准偏差或相对偏差的检验

截距检验也可以采用剩余标准偏差或相对偏差的检验，也可以根据经验值，其容许范围对于不同的分析方法和精密度要求，可做适当调整。

#### 6.2.2.3.3 使用质控图

根据专业技术领域可接受范围，指定截距的警戒限和工作限，绘制质控图（参见6.5）。

#### 6.2.2.4 斜率检验

校准曲线的斜率常随环境温度、试剂批号和贮存时间等实验条件的改变而变动，需要在分析待测样本的同时分析校准曲线的系列浓度校准样本，使用剩余标准偏差S，或相对偏差进行统计分析。

#### 6.2.2.5 校准曲线核查

包括不同系列浓度校正标样和中间浓度校正标样核查两种方式。

同系列浓度校正标样和中间浓度校正标样应采用不同来源的标准储备液或者储备物质制备，以避免出现与方法分析性能无关的偏差，并且校正曲线至少有独立两条，包含至少5个不同浓度水平的校正标样。

核查浓度与标示值的偏差应该满足方法标准或者专业技术领域的要求，如果不满足要求，按照规定处理。例如生物样本中药物及代谢物浓度测定，校准曲线的不同系列浓度的校准标样浓度的回算浓度应在定量下限标示浓度的 $\pm 20\%$ 以内，其他水平应在标示值的 $\pm 15\%$ 以内。至少75%的校准标样符合要求，如果不符合接受标准，则应拒绝该校正标样，并重新拟合去除该点后的校准曲线。

### 6.3 加标回收

#### 6.3.1 实施要点

加标回收是实验室一个重要的经常使用质控的手段，能够反映分析全过程的总体质量水平，通过测定加标回收率评价分析过程的质量。一般加标回收应考虑以下要素（包括但不限于）：

- 选择符合要求的空白样品或者待测样品，不与待测物发生反应，不干扰待测物的准确测定；
- 加标用标准物质或标准溶液来源明确，具有准确量值，应该满足分析方法的要求，必要时要进行核查；
- 加标量应确保准确，使用经过校准/检定并确认的计量器具；
- 在空白样品或者待测样品中准确加入标准物质或者标准溶液，制备成均匀的加标样品；
- 加标量一般需和待测物浓度相近，在待测物浓度极低时，需按检测下限/定量下限的量加标，如果待测样品浓度范围较宽，加标样品应该设置不同加标量，应该覆盖待测样品浓度范围，不得超过待测物浓度的2倍；
- 加标回收样品数量需根据本专业技术领域以及技术指导原则的要求，在分析批中分析符合要求的数量，例如血浆中药物浓度测定，不得少于分析批样本数的5%。

#### 6.3.2 结果评价

加标回收的结果，通常按加标回收率进行评价，回收率的可接受限包括方法标准的回收率范围、专

业技术领域以及技术指导原则的回收率可接受范围以及根据加标量或者待测物含量（常量、微量、痕量）情况满足相应回收率的要求。当使用控制图方法时，根据专业技术领域可接受范围，指定回收率的警戒限和工作限，绘制质控图（参见6.5）。

加标回收率计算公式：

$$\text{回收率} = (\text{加标样品测定值} - \text{空白样品或待测样品的测定值}) / \text{加标量} \times 100\% \cdots \cdots (1)$$

## 6.4 空白样品

### 6.4.1 使用方法

空白样品一般与样品“同流程”分析并在待测样品分析前进行分析，可用于（包括但不限于）：

- 评估分析系统的适用性；
- 监控分析系统的污染情况；
- 检查待测物或内标的保留时间处没有干扰组分引起的显著响应。
- 空白样品应该来源明确，容易获得
- 基质均匀性好，在待测样品同样储存和处理条件下稳定
- 基质应与待测样品一致，不含待测物，不与待测物发生反应。在难以获得与待测样品基质一致的空白样品时，可使用合适的替代基质作为空白样品。

### 6.4.2 结果评价

空白样品的评价包含但不限于以下方式：

- 依照方法标准要求；
- 专业技术领域以及技术指导原则要求；
- 使用控制图方法，根据专业技术领域可接受范围，指定不要超出检出限或者定量限响应值的警戒限和工作限，绘制质控图（参见6.5）；
- 如果空白样品存在干扰组分，干扰组分的响应应满足要求：
  - 干扰低于分析方法的检出限，视为无干扰；
  - 干扰高于分析方法的检出限，量值恒定时，可按差减法处理；
  - 根据专业技术领域的相关要求处理，例如生物样本中药物浓度测定，干扰组分应不高于待测物定量下限响应的20%，并且不高于每个基质定量下限样品中内标响应的5%。

## 6.5 控制图

### 6.5.1 作用

控制图是内部质量监控的一个重要工具，它用来持续监控和评估一指定分析测试系统的稳定性、精密度和偏差。

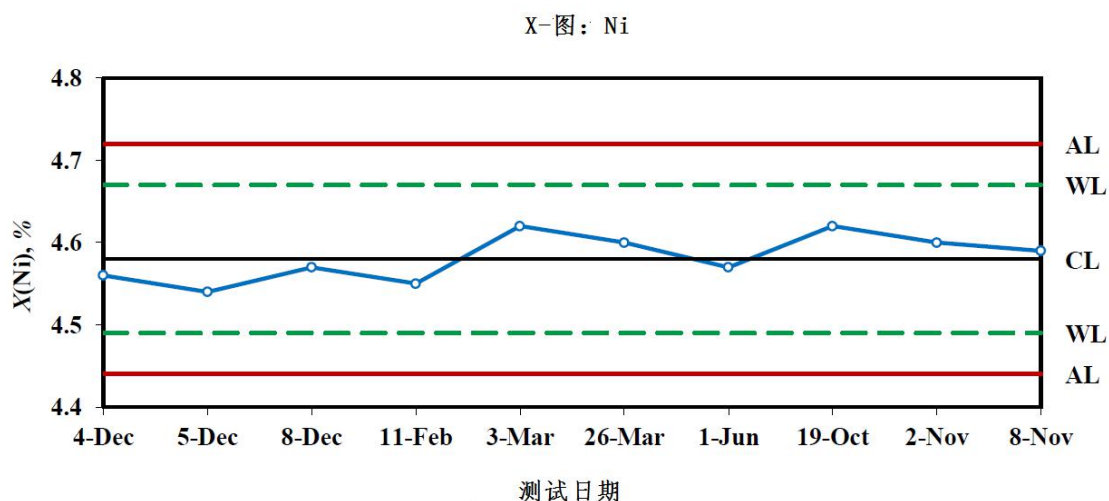
一个分析测试系统，是指包括从抽样、实验室测试直到结果报告的过程，因此，一个标准测试方法就是一个分析测试系统。

控制图的制作和应用是基于这样的假设：即处于统计受控状态的测试系统其行为呈正态分布。

典型的控制图见图XXX，通常包括一条中位线(central line, CL)和四条控制限。中位线相当于控制样测试结果的平均值。四条控制限，分别为上、下警告限(warning limit, WL)和上、下行动限(action limit, AL)。上、下警告限分别位于中位线两侧并距离中位线两倍标准偏差(2s)的位置，意味着呈正态分布的控制样数据约95%落在其间。而上、下行动限于中位线两侧并距离中位线三倍标准偏差(3s)的位置，意味着呈正态分布的数据约99.7%落在其间。

在控制图中，如果所有控制样测试值都落在上下警告限之间，表明分析测试系统在规定的限值范围内运行，可以报告待测样品的分析结果。如果控制样测试值落在上下行动限之外则说明分析测试系统有问题，不得报告待测样品的分析结果，而应采取纠正行动，识别误差的来源并予以消除。如果控制值落在警告限之外但在行动限之内，则应根据特定的规则进行评估(详见6.5.6 I控制图的解释及结论的使用)。

典型的控制图如下：



常用的控制图包括I图(单值图)、X-图(单值图或均值图)和R-图(极差图)，其他还包括MR-控制图、EWMA-控制图。

### 6.5.2 控制样品选择

控制样必须是稳定、均匀，其基质应与待测样品尽可能相同，并且其量应足以保证日常测试的质量监控以及为建立控制图控制限所需进行的测试所使用。控制样可以是有证标准物质、实验室内部质控样品以及满足上述条件的其他物质。

### 6.5.3 控制样数据的准备

为了制作控制图，需要获得与控制样数据正态分布有关的参数，如标准偏差。可以采用历史数据，如没历史数据，则要通过测试至少获得20个数据。由于控制图是监控测试系统的长期稳定性，测试数据的标准偏差需要反映实验室在不同时间，由一个人或多个人使用同一个测试方法，使用同类型的不同仪器，对同一个控制样进行测试的情况。因此，实验室应在期间测量精密度测量条件下来获得控制样的测试数据，对于初次建立的控制图，如没有历史数据时，实验室应将实验室内部质控样或有证标准物质与正常测试的一组样品一起进行测试，可行时，在正常测试段内随机插入到样品的检测序列中，为反映期间精密度的条件，这20个数据应至少分散到15天内获得，此外，在正常测试时如需要使用不同的仪器和由不同人员来测试时，则在获取控制样数据时也需要使用不同的仪器，并由不同的人员来测试。在该条件下所获得的标准偏差(即期间测量精密度)用 $S_R$ 来表示。

在日常测试时，将控制样与正常测试样品一起测试，可行时，随机插入到样品的检测序列中。适用时，一个分析测试批至少有一个控制样。一个分析测试批的控制样数量取决于测试系统的稳定性以及控制样数据超出行动限而造成整个分析测试批需要重新测试的成本有关。

注：不能在重复性条件下获得这些数据，否则利用这些数据计算所得的标准偏差会被低估。

### 6.5.4 数据的处理

按6.5.3获取20个数据后，在XY坐标图上，以检测次数（或检测日期）为横轴，将所测得的控制样品数据（对于有证标准物质，为控制样测试结果减去其值后的数据）按检测顺序（或日期）依次标在纵轴上，形成一个控制样数据随检测次数（或日期）走向的趋势图。把这些点用线连起来，以便更清楚地观察数据的走向。X轴的空间要足够大以便容入不断加入的数据，Y轴的空间也要考虑到数据的最大值和最小值。

目测检查是否有可疑的数据，如由转录误差引起的数据就与其他数据明显不同，剔除这类数据，剔除后如数据不够15个，再补充数据。

用检验离群值的方法进行检验，如有，剔除相关的数据，见附录XX 检验离群值的方法

再检查是否有异常排列，即与随机分布应有的排列不一致的数据，如连续一边倒（连续上升或下降）、非正常聚堆和循环，如有，则不用这组数据，重新准备数据。

对数据进行正态性检验。数据正态性是制作控制图的基础，非正态分布的数据不能用来制作控制图。如以往对这类数据已检验过正态性并符合要求，则不需要再做正态性检验。正态检验可用画正态概率图和the Anderson-Darling 统计量。正态检验合格后开始制作控制图。

正态性检验具体程序见GB/T 27407-2010

## 6.5.5 控制图的制作

### 6.5.5.1 I-图

在按6.5.4处理后由合格控制样数据形成的趋势图上，添加中位线和控制限后就变成I-控制图，I-控制图是一个单值控制图，可用于监控控制值的系统效应和随机效应。当使用实验室内部质控样时可监控测试系统的精密度，不能监控正确度，而使用有证标准物质时则可以监控两者。

#### 6.5.5.1.1 中位线 CL

设实验室内部质控样或有证标准物质的检测结果单值为 $X_i$ ，而将其用于制作I图的数据为 $Y_i$ ， $Y_i$ 与 $X_i$ 的关系如下：

- (1) 使用实验室内部质控样时

$$Y_i = X_i \quad \dots\dots \quad \text{公式 (6)}$$

- (2) 使用有证标准物质时

设有证标准物质的值为 $\mu_0$ ，则：

$$Y_i = X_i - \mu_0 \quad \dots\dots \quad \text{公式 (7)}$$

质控样（包括标准物质）累积数据的算术平均值 $\bar{Y}$ ，

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} \quad \dots\dots\dots \quad \text{公式 (8)}$$

将该 $\bar{Y}$ 数值标在XY坐标图的纵坐标上，然后通过该点画一条与横坐标平行的直线，该线即为中位线 CL。

#### 6.5.5.1.2 统计控制限

- (1) 期间测量精密度标准偏差估计值 $S_{R'}$ ：

$$S_{R'} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1}} \quad \dots\dots\dots \quad \text{公式 (9)}$$

式中：

$i$  —— 累积的检测结果数据的个数， $i = 1, 2, \dots, n$ ；

$n$  —— 累积的检测结果数据总数；

$Y_i$  ——质控样的单值检测结果数据或其减去标准值后的数据；

$\bar{Y}$  ——累积检测结果数据或其减去标准值后数据的平均值。

(2) 行动限AL

$$AL = \bar{Y} \pm 3 \times S_R \quad \dots\dots\dots \text{公式(10)}$$

(3) 警告限WL

$$WL = \bar{Y} \pm 2 \times S_R \quad \dots\dots\dots \text{公式(11)}$$

将该AL与WL的数值标在XY坐标图的纵坐标上，然后通过这些点分别画一条与横坐标平行的直线，这些线分别为代表行动限和警告限的线。见附录AXXX 用EXCEL制作控制图。

### 6.5.6 I 控制图的解释及结论的使用

要考虑两种情况：

1) 数据超出上行动限或下行动限

个别数据如超出上行动限或下行动限，这是测试系统不受控的强烈明确信号具体表述，需要调查其根本原因，在根本原因被调查或被纠正之前，同时测试的其他结果应被列为可疑。这时可采取以下步骤：

当出现“不受控”时，通常再对控制样测试至少2次以上，如控制样的新测试结果位于警告限内，则可对该同批测试的样品再重新进行测试，如果控制样新的测试结果处在警告限外，则不应对该同批测试的样品再重新测试，因为测试系统可能存在误差，要采取纠正措施，找出原因并加以消除后才可以再对该批样品进行测试。进行纠正措施时，可从人机料法环各方面对测试过程存在的问题进行排查，如排除后，对控制样重新测试的结果仍然是超出行动限，则不能报告该同批测试的样品的测试结果。

2) 数据未超出上行动限或下行动限

数据如出现下列情况，测试系统可能出现了“不受控”的状况：

还有其他情况上中位线同一侧连续3个点中的2点落在 $2S_R$ （2个标准偏差）之外；

b) 中位线同一侧连续5个点大于 $1S_R$ （1个标准偏差）；

c) 连续9点或更多点落在中位线的同一侧；

d) 连续7个点稳步递增或递减。

当出现上述情况时，可认为测试系统依然受控，可报告同批测试的样品测试结果，但存在潜在失控的可能，需要调查原因，找到问题所在，再进行样品测试。

### 6.5.7 更新 I 控制图的参数

两种情况下需要对控制图的参数进行更新。

(1) 新建立以及运行时间较短的控制图

新建立的控制图，由于其数据是在比较短的时间内获得，其标准偏差尚不能完全反映实验室的期间测量精密度条件，因此，要随着控制样数据的增多，定期更新标准偏差及其控制限的数据，此过程持续进行，直到用来计算标准偏差的数据达到100个，此时，实验室可以将控制图的参数固定下来了，也可以继续更新。更新方法如下：

当收集至少20个新数据后，可用它更新旧数据组以得到新的精密度估计值。在更新前，先进行F检验（见附录A），将新数据组的样本方差与旧数据组的样本方差进行比较。假如，F检验没有发现显著差异，则合并两者的样本方差得到更新后精密度。如旧的样本方差是基于超过100个数据所得的，则可以不更新精密度。如有显著差异，需要调查根本原因。

合并样本标准偏差计算公式如下：

$$S_{\text{合并}} = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \quad \dots\dots\dots \text{公式(12)}$$

式中：

$n_1$ : 控制图上旧数据组的数据数量;

$n_2$ : 控制图上新数据组的数据数量;

$S_1$  现有控制图所使用的标准偏差;

$S_2$  新数据的标准偏差。

中位线为新旧质控样的所有数据的平均值, 即:

$$\bar{Y} = \frac{\sum_i^n Y_i}{n} \dots\dots\dots \text{公式 (13)}$$

其中:  $n = n_1 + n_2$ ;

更新后, 用新的标准偏差 (即 $S_{\text{合并}}$ ) 和中位线更新警告限、行动限。

## (2) 用新质控样时

当现有质控样快用完时, 如其量不足以支持超过20个数据的测试时, 就应开始准备新的质控样, 并将之与旧质控样同时测试。每次均同时记录新旧质控样的测试记录, 只当旧质控样的数据在控制图上显示正常, 也即测试系统是受控时, 同时所测的新的质控样的数据才有效。当累积至少20个新的质控样的受控数据后, 进行F检验 (见附录A), 将新数据组的样本方差与相同浓度的历史数据组的样本方差进行比较。假如, F检验没有发现显著差异, 则按上节的方法合并两者的样本方差得到更新后精密度。并更新中位线, 中位线的值按上节方法进行计算。如历史数据的样本方差是基于超过100个数据所得的, 则可以不更新。如有显著差异, 需要调查根本原因。

## 6.5.8 其他控制图

其他控制图: 如X-图、R-图 (或 $r\%$ -图)、MR-图、EWMA-图等, 具体见GB/T 27407。

1) X-图: 为平均值-标准偏差的控制图, 用于监控系统因素和随机因素对控制值产生的联合效应。

(2) R-图: 即极差-图, 是指在重复性条件下多次检测同一个样品, 计算每次重复性条件下的极差(R), 然后利用这些极差(R)数据所建立的控制图, 其中极差(R)是指两个或两个以上独立样品的单个测量结果中最大值和最小值之差。

如采用相对极差 ( $r\%$ ), 即极差与样品浓度 (在检出限水平以上) 的比, 得到的控制图即为 $r\%$ -图。

R-图的首要目的是监控重复性。最简单的R图只进行双样重复测试, 即在一个分析批中对待测样品进行双样重复测试, 计算两个平行结果之间的差值, 然后将差值绘制在控制图上, 则可得到最简单的R-图。

(3) MR-图: 是移动差-图 (moving range chart), 也叫移动极差-图, 指将按时间顺序获得的两个相邻数据相减并取绝对值以得到MR值, 再利用这些MR值所建立的控制图。用于监控测试系统总方差的变动。适用于质控样给出单值的情况。

(4) EWMA-图: 指数加权移动平均值 (EWMA, 即exponentially weighted moving average)-图, 指将现有数据与以前的结果加权平均, 其中权重随着数据离现有数据的时间而呈指数级下降, 以得到EWMA值, 利用EWMA值建立的控制图。通常将EWMA趋势线叠加到I-图上, 以提高检测平均值漂移的灵敏度。适用于质控样给出单值的情况。

具体见GB/T 27407、GB/T 32464

## 6.6 设备期间核查

### 6.6.1 目的

设备 (包括测量设备和标准物质) 应进行期间核查, 以保持对设备性能的信心。实验室的设备期间核查程序应包括期间核查的范围 (具体设备)、核查方法要求以及实施要求等。根据设备性能的特点和核查标准的易获得性, 通过期间核查一般可实现:



- a) 确认设备的外观完好性、功能正常性；
- b) 确认设备校准状态的可信性；
- c) 验证设备的性能是否持续符合规定要求。

## 6.6.2 对象的选择

需要进行期间核查的对象包括：

- a) 使用年限较长，其计量性能的稳定性和可靠性下降，即稳定性变差、漂移较大的测量设备；
- b) 使用频率高的测量设备；
- c) 经常携带到现场进行检测以及在恶劣环境下使用的测量设备；
- d) 测量重要、关键项目数据和测量准确度要求较高的测量设备；
- e) 性状不稳定、储存条件发生变化、使用频率高以及测量结果不可信的在有效期内的标准物质。

## 6.6.3 方法

### 6.6.3.1 使用计量校准

具有准确度（准确度等级或最大允许误差）要求的检测仪器设备，可利用计量标准对其期间核查，要求计量标准的最大允许误差要求小于核查设备的最大允许误差的1/3。

### 6.6.3.2 使用标准物质

#### a) 有证标准物质

具有能溯源到国际单位制（SI）的有证标准物质时，可利用有证标准物质对其检测所使用的仪器设备进行期间核查，要求核查设备的示值与标准物质定值之差小于等于核查设备的最大允许误差。

#### b) 其他标准物质

实验室也可以根据实际工作需要选定合适的核查标准，用于对其检测使用的仪器设备进行期间核查。但实验室要能够确保选定的核查标准性能稳定且不做其他用途。如，使用质控样品复测或留样复测。

### 6.6.3.3 使用设备比对

#### a) 两台设备比对

可以溯源到同一计量标准两台同类型测量设备，且其中一台设备检定/校准后能确定其检定/校准状态的可信度，可采用两台设备比对的方法，同时对两台测量设备进行期间核查

#### b) 多台设备比对

当实验室没有高一等级的测量设备来采用直接测量法，但有多台具有相同测量准确度等级（确定度等级、最大允许误差或测量不确定度）的设备，可以采用多台设备比对的方法，利用比对测量设备测量结果的一致性，同时对多台测量设备进行期间核查。

### 6.6.3.4 其他方法

仪器使用说明书和检测方法对使用的化学检测仪器设备有特定要求的，要按照使用说明书或者检测方法的要求对该化学检测设备进行期间核查，确认设备满足使用要求。

可用质控图（详见6.5）。

## 6.6.4 期间核查的结果评价和处理

结果评价方法可以参考相关章节或描述 多台设备比对没处引用

实验室通过期间核查发现测量设备的状态接近实验室核查判定标准时,要查找原因并采取适当的风险防控措施,例如增加核查频次。实验室通过期间核查发现测量设备的状态超出实验室核查判定标准时,需停用该设备,并依照实验室相关的仪器管理程序实施。

### 6.6.5 期间核查的文件化

实验室宜制定化学检测仪器设备期间核查文件,可包括以下内容(针对不同类别的设备选择适当期间核查文件体系):

- a) 适用范围;
- b) 制定依据;
- c) 核查的环境要求;
- d) 核查使用的设备和计量标准、标准物质或核查标准;
- e) 核查内容;
- f) 核查方法;
- g) 判定准则;
- h) 操作步骤;
- i) 结果评价和结论。

## 6.7 留样复测

### 6.7.1 用途

在不同的时间(或合理的时间间隔内)再次对已完成检测的留存样品或备份样品进行检测。可用于以下情形:

- 对留存样品特性的监控;
- 验证测试结果的公正性;
- 验证测试结果的复现性。

### 6.7.2 样品选择

留样复测宜选择至少有一次测量结果、所检测项目的含量应在检测方法定量限以上且在保存期限内性能稳定。

### 6.7.3 结果评价

6.7.3.1 使用标准方法的,用两次检测结果的测量不确定度或规定的再现性限进行评价。在没有测量不确定度或再现性限时,可使用两次检测结果的绝对差值不大于平行试验的允许差进行评价。

6.7.3.2 利用多个留样样本复测的数据,求得中间精密度 $\sigma_1$ 后,前后两次结果之差不能大于 $2.8\sigma_1$ ,其中 $\sigma_1$ 为中间精密度标准偏差:

$$S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n}} \quad \dots\dots\dots \text{公式 (14)}$$

式中:

- $S_d$  —— 差值的标准差;
- $d_i$  —— 每个差值
- $\bar{d}$  —— 差值的平均值
- $n$  —— 差值的数量

$$\sigma_1 = \frac{S_d}{\sqrt{2}} \quad \dots\dots\dots \text{公式 (15)}$$

6.7.3.3 按照行业相关规则进行评价。

## 6.8 重复分析

### 6.8.1 重复分析用途

重复分析一般指在重复性条件下实施的检测活动，主要用于核查或评估检测系统/检测过程的精密程度水平。当重复分析样品为2个时，也称平行分析。

### 6.8.2 重复分析安排

重复分析样品为同一样品的两份或多份子样。重复分析样品的数量可根据样品的批量、测定的难易程度等因素确定。实验室可制定重复分析样品的抽取规则，随机抽取或按规定顺序抽取。重复分析样品一般由同一人分析，也可以交由不同人员分析，有条件时按盲样分析，数据更为客观。当标准方法规定了重复分析样品数量，按标准方法规定执行。如标准方法没有规定，可参考如下方式（包括但不限于）：

- a) 一般不少于全部样品或一个分析批次的 10 %；
- b) 特殊样品或质量要求较高的样品可酌情增加数量直至 100 %；
- c) 当一批样品的数量较少时，至少安排一个重复分析样品，也可安排 100%的重复分析。

### 6.8.3 重复分析结果评价

实验室可制定重复分析结果的可接受限，做为评价依据，一般可采用如下方法（不限于）：

- a) 依据与客户的约定或相关标准规定评价；
- b) 依据使用标准方法的重复性限（ $r$ ）作为重复分析结果的可接受限；重复分析结果之差的绝对值小于或等于  $r$  时为满意结果；大于  $r$  时为不满意结果；
- c) 采用重复分析样品结果的相对偏差（ $RD$ ）按照规定的可接受限评价，按公式（16）或（17）计算相对偏差：

$$RD = [ (X_i - \bar{X}) / \bar{X} ] \times 100 \% \quad (i=1,2,\dots,n) \quad (16)$$

当  $i=2$  时：

$$RD = [ (X_1 - X_2) / ( (X_1 + X_2) / 2 ) ] \times 100 \% \quad (17)$$

式中：

$RD$  为相对偏差；

$X_i$  为检测结果  $i$  ；

$X_1$  为检测结果 1 ；

$X_2$  为检测结果 2 ；

$\bar{X}$  为  $i$  个检测结果的平均值。

- d) 当已知重复分析的标准偏差(S)或进行多份重复样品分析时,可按公式(18)计算相对标准偏差(RSD),依据规定的可接受限进行评价:

$$RSD = ( S / \bar{X} ) \times 100 \% \quad (18)$$

式中:

$RSD$  为相对标准偏差;

$S$  为标准偏差;

$\bar{X}$  为  $n$  个检测结果的平均值(一般  $n > 5$ )。

- e) 当已知或可评定所使用方法结果的测量不确定度时,可按公式(19)评价平行分析结果。比值小于等于1时,满意结果,反之为不满意结果。

$$\text{即当 } \frac{|X_1 - X_2|}{\sqrt{2}U} \leq 1 \text{ 时, 为满意结果。} \quad (19)$$

式中:

$X_1$  为检测结果1;

$X_2$  为检测结果2;

$U$  为检测结果的扩展不确定度( $p=95\%$ ,  $k=2$ )。

## 6.9 盲样测试

### 6.9.1 用途

由实验室质控管理人员进行组织和设计盲样测试样品和盲样考核方式,将盲样以密码方式发放给检测人员。

### 6.9.2 样品选择

盲样测试样品一般选择有证标准物质、标准物质和质量监控样品。

### 6.9.3 结果评价

见本文件 6.1.3。

## 6.10 仪器比对

### 6.10.1 适用情形

实验室可采用仪器比对试验的方式进行内部质量监控,判断对测量准确度、有效性有影响的仪器设备是否符合测量溯源性的要求,用以评价仪器设备对实验室检测结果准确性、稳定性和可靠性的影响。适用于原理相同的仪器间比对,如相同厂家的相同型号或不同型号仪器比对,或不同厂家生产的同类仪器比对,根据结果的差异来评价是否满足要求。

作为内部质量监控手段,仪器比对试验优先适用于以下情况:

- a) 新安装的仪器设备;
- b) 修复后的仪器设备;
- c) 检测结果出现在临界值附近的仪器设备。

## 6.10.2 结果评价

### 6.10.2.1 两台仪器比对

当一台设备与另一台设备进行比对时，将已知准确度等级设备的测定值作为参考值，比对试验结果按式（20）进行评价：

$$|y_1 - y_0| \leq \sqrt{U_1^2 + U_0^2} \dots\dots\dots (20)$$

式中：

$y_1$ ——仪器A的测定值；

$y_0$ ——仪器B的测定值；

$U_1$ ——比对方测定值的测量不确定度， $k=2$ ；

$U_0$ ——参考方测定值的测量不确定度， $k=2$ 。

若满足式（5-10-1），表明比对试验结果满意；若不满足式（21），表明比对试验结果不满意。

### 6.10.2.2 多台仪器比对

当对多台相同准确度等级的设备结果按式（5-10-2）进行评价。

$$|y_i - \bar{y}| \leq \sqrt{\frac{n-1}{n}} \times U \dots\dots\dots (21)$$

式中：

$y_i$ ——第*i*台设备的测定值；

$\bar{y}$ ——多台相同准确度等级设备测定值的平均值；

$n$ ——参加比对的设备台数；

$U$ ——参加比对的多台相同准确度等级设备的测量不确定度， $k=2$ 。

若满足式（5-10-2），表明比对实验结果满意；若不满足式（21），表明比对试验结果不满意。

## 6.11 方法比对

### 6.11.1 适用情形

当某个检测项目可由多种检测方法进行操作时，实验室可采用方法比对试验的方式进行内部质量监控，判断检测所遵循的标准或方法是否被正确理解并严格执行，用以评价检测方法对实验室检测结果准确性、稳定性和可靠性的影响。作为内部质量监控手段，方法比对试验优先适用于以下情况：

- a) 新标准、新方法或自建的新方法；
- b) 有多个检测标准或方法的项目。

### 6.11.2 结果评价

#### 6.11.2.1 方法规定了重复性限或再现性限时，比对试验结果按以下方法进行评价：

在对重复性条件或再现性条件下得到的两个单一测试结果进行试验结果比对时，应与重复性限  $r = 2.8\delta_r$  或再现性限  $R = 2.8\delta_R$  进行比较。两个方法测试结果之差的绝对值应用  $r$  或  $R$  来检验，如果绝对差小于或等于  $r$  或  $R$ ，即认为结果一致，如果两个测试结果的绝对差大于  $r$  或  $R$ ，必须找出差异的原因是否由于测试方法的精确度和测试样本有偏差。

注：应该注意到， $\delta_r$  和  $\delta_R$  的真值往往是不能得到的，所以在实际的统计计算中是通过精密度试验，用所有参加实验室的结果经统计计算得到的重复性标准差和再现性标准差估计值（分别表示为  $S_r$  和  $S_R$ ）来代替。

6.11.2.2 方法未规定重复性限或再现性限时，方法比对试验结果按式（22）进行评价：

$$|y_1 - y_2| \leq \sqrt{U_1^2 + U_2^2} \dots\dots\dots (22)$$

式中：

$y_1$ ——第1种方法的测定值；

$y_2$ ——第2种方法的测定值；

$U_1$ ——第1种方法的测定值的测量不确定度， $k=2$ ；

$U_2$ ——第2种方法的测定值的测量不确定度， $k=2$ 。

若满足式（22），表明比对试验结果满意；若不满足式（22），表明比对试验结果不满意。

## 6.12 人员比对

### 6.12.1 适用情形

实验室可采用人员比对试验的方式进行内部质量监控，通过安排具有代表性的不同层次的两人或多人开展比对试验，考核检测人员能力水平，判断检测人员操作是否正确、熟练，用以评价人员对实验检测结果准确性、稳定性和可靠性的影响。作为内部质量监控手段，人员比对试验一般适用于以下情况：

- 依靠检测人员主观判断较多的项目，如食品中感官、品尝等项目；
- 在培训的员工和新上岗的人员；
- 检测过程的关键控制点或关键控制环节；
- 操作难度大的样品或检测项目；
- 检测结果在临界值附近；
- 新开展的检测方法。

### 6.12.2 结果评价

6.12.2.1 方法规定了重复性限或再现性限时，人员比对试验结果还可按以下方法进行评价：

在对重复性条件或再现性条件下得到的两个单一测试结果进行试验结果比对时，应与重复性限  $r = 2.8\delta_r$  或再现性限  $R = 2.8\delta_R$  进行比较。

两个方法测试结果之差的绝对值应用  $r$  或  $R$  来检验，如果绝对差小于或等于  $r$  或  $R$ ，即认为结果一致，如果两个测试结果的绝对差大于  $r$  或  $R$ ，必须找出差异的原因是否由于测试方法的精确度和测试样本有偏差。

注：应该注意到， $\delta_r$ 和 $\delta_R$ 的真值往往是不能得到的，所以在实际的统计计算中是通过精密度试验，用所有参加实验室的结果经统计计算得到的重复性标准差和再现性标准差估计值（分别表示为 $S_r$ 和 $S_R$ ）来代替。

6.12.2.2 方法规定了允许差时，进行人员比对的测定值作为参考值，比较按相同标准方法测定相同项目的试验结果，比对试验结果按式（23）进行评价：

$$\frac{|x_i - x|}{x} \times 100 \leq D\% \dots\dots\dots (23)$$

式中：

$x_i$ ——比对方的测定值；

$x$ ——参考方的测定值；

$D\%$ ——方法规定的允许差，以百分数（%）计。

若满足式（23），表明比对试验结果满意；若不满足式（23），表明比对试验结果不满意。

6.12.2.3 方法没有规定允许差时，进行人员比对的双方应对测定值的测量不确定度进行评定，比对试验结果按式（5-12-2）进行评价：

$$E_n = \frac{|y_1 - y_2|}{\sqrt{U_1^2 + U_2^2}} \dots \dots \dots (24)$$

式中：

$y_1$ ——比对方人员的测定值；

$y_2$ ——比对方另一人员的测定值；

$U_1$ ——比对方人员测定值的测量不确定度， $k=2$ ；

$U_2$ ——比对方另一人员测定值的测量不确定度， $k=2$ 。

当 $E_n \leq 1$ 时，表明比对试验结果满意；当 $E_n > 1$ 时，表明比对试验结果不满意。

当结果不满意时，实验室需要分析原因，采取纠正措施进行整改。

## 6.13 结果相关性

### 6.13.1 综述

分析物品不同特性结果的相关性是确保实验室检测结果有效性的监控方式之一，通过分析同一物品不同特性参数之间的关联性质，以相关联检测结果的特定逻辑关系为判定依据，对一组检测结果的合理性做评估，有助于保证检测结果的准确性及有效性，提高工作效率。在化学检测中应用广泛，通常分为以下几类情况。

6.13.2 包含关系，也即总量大于分量，很多经典的筛查（screening）检测方式就是利用不同特性结果之间的包含关系，在保证检测结果准确性的前提下，提高了检测效率。例如：

——总量金属大于等于溶解性金属；

——固体总含量大于浸出含量；

——含量与有效态的关系，比如总铬含量大于六价铬含量；

——全消解的含量应比半消解的高；

——食品中总膳食纤维值大于其中的可溶性膳食纤维及不可溶性膳食纤维。

6.13.3 不同参数之间有相关性，可拟合出较为明确、甚至是可量化的函数关系，举例如下：

#### 1) 水中阴阳离子平衡的关系

阴离子物质的量浓度之和（ $\sum Ca$  mmol/L）与阳离子的物质的量浓度之和（ $\sum Cc$  mmol/L）是否接近，以及其最大允许差（R）： $R = [(\sum Ca - \sum Cc) / (\sum Ca + \sum Cc)] \times 100\%$ ，作为判断结果是否合理的依据：

天然水，当 $\sum Ca + \sum Cc > 5$  mmol/L 时， $R \leq \pm 3\%$ ；

污染水，当 $\sum Ca + \sum Cc < 5$  mmol/L 时，可不考虑。

#### 2) 水中可溶性固体总量与阴阳离子总和的关系

用蒸干法测得的可溶性固体总量（A），与实测的各组分含量的总和减去重碳酸根离子含量的一半所得的结果（B）是否接近，其最大允许相对误差（Er）： $Er = [(A - B) / (A + B)] \times 100\%$ ，作为判断结果是否合理的依据：

当 $A < 100$  mg/L 时， $Er \leq \pm 5\%$ ；

当 $A > 100$  mg/L 时， $Er \leq \pm 3\%$ 。

3) 岩矿主成分分析时，物质中各组分的加和应等于 100%，岩石、矿石、矿物全分析各组分

除按重复分析相对偏差允许限检查外，其主要组分各项结果的百分数加和可分两级检查：

第 1 级：99.3 %—100.7 %

第 2 级：99.0 %—101.0 %

各项百分数加和的检查级别，依据试样的特性和客户的要求确定。

4) 总氮=总凯氏氮+硝酸盐氮+亚硝酸盐氮；凯氏氮=有机氮+氨氮。

#### 6.13.4 不同参数之间有相关性趋势，但不能拟合出可量化的函数关系，举例如下：

1) 如水样的浊度和悬浮物的关系，悬浮物颗粒越多，浊度越大；

2) 纺织品中如检出禁用物质-四氨基偶氮苯，样品中很可能还会含某几类分散染料（如分散黄 3）。

样品感官性状，如颜色、浑浊和气味等，样品的颜色深浅，与色度测试结果应大致一致；气味与一些测试结果相关，如样品有臭味，可能硫化物或有机物（COD、凯氏氮、氨氮）含量较高。

## 7 外部控制方法

### 7.1 能力验证

能力验证（含测量审核），是对单个实验室或一组实验室进行的周期性评估，是对分析系统误差进行的间歇性检查，是证明实验室检测结果准确可靠的重要手段。实验室可通过参加能力验证，辅助建立有效的质量监控系统。参加能力验证不仅是实验室外部质量监控的重要手段，同时也是内部质量监控的补充。作为外部质量监控手段，参加能力验证应按以下原则进行：

- a) 参加实验室能力验证计划是实验室质量监控系统有效运行、检测结果准确可靠的间接证明，但不能代替实验室内部质量监控。
- b) 实验室应根据检测能力范围，合理安排参加实验室能力验证的计划，实验室参加能力验证活动领域和频次应满足相关规定要求。

注：能力验证是对单一实验室或一组实验室进行的周期性评估，参加能力验证有助于辅助建立有效的质量监控系统，但是并不能代替实施内部质量监控措施，反之亦然。如果没有良好的内部质量监控系统支持，实验室参加一次能力验证所获得的评价并不能说明问题。

### 7.2 测量审核

#### 7.2.1 测量审核类型与评价方法

测量审核 (Measurement Audits)，是实验室对被测物品 (材料或制品) 进行实际测试，将测试结果与参考值进行比较的活动。测量审核是能力验证计划的一种，有时也称为“一对一”的能力验证计划。参加者若选择测量审核，需核查测量审核样品来源、稳定性、指定值和能力评定标准差的确定方式，是否符合本实验室既定项目的质量监控的目标。根据测量审核样品的来源，通常有以下几种类型：

##### 7.2.1.1 采用能力验证样品

对于采用已开展过能力验证项目的样品，通过用实验室的测定结果与该项目指定值和标准偏差进行比对，以统计的方法计算稳健Z比分数来评价实验室的检测能力。

##### 7.2.1.2 采用有证参考物质



对于采用有证参考物质，通过用实验室的测定结果与该参考物质的参考值进行比对，而后根据与实验室采用国际标准、国家标准或精度相当的方法的精密度数据来评价参加实验室的检测能力，采用 $E_n$ 值判断，或者按照临界值（CD 值）评定。

### 7.2.1.3 采用参考物质（样品）

当用于测量审核的参考物质（样品）无指定参考值时，应指定参考实验室采用与申请实验室相同的测试方法进行测试，而后用申请实验室的测定结果与参考实验室的测定结果进行比对，根据与实验室采用检测方法相当精度的国际或国家标准方法的精密度或允许差来评价参加实验室的检测能力，具体评价方法与采用有证参考物质（样品）时的评价方法相同。

### 7.2.1.4 定性或半定量项目

对于定性或半定量性质的测量审核项目，测量结果的评价按实验室所采用相当精度国际、国家标准的规定执行。

## 7.2.2 测量审核的选择、核查、评价与利用

当测量审核样品为能力验证剩余样品时，参加者需核查实施机构是否确认样品稳定，能力评定方式是否合理。如样品被测特性不稳定，参加者可要求实施机构需重新定值或重新选用特性值稳定可靠的样品。能力评定标准差可采用经验值，或上一轮能力验证计划确定的能力评定标准差，通常采用稳健 $Z$ 比分数评价方法评定参加者的能力。

当测量审核样品为标准物质/标准样品或质控样品，通常参照本标准第5.1.4章的统计方法对参加者的检测结果进行评价。

当测量审核样品为定制样品时，参加者需核查实施机构是否采用“指定值的确定”中的适配方法或其他合理方式确定指定值，能力评定标准差可采用经验值或其他合理的方式确定，比如利用Horwitz 公式或方法精密度确定，或者采用稳健 $Z$ 比分数评价方法评定参加者的能力。

## 7.3 实验室间比对

### 7.3.1 比对方案的策划

比对方案适用于两个及以上实验室间比对，其目的是为了确定实验室所使用方法的可靠性和检测结果的准确性。方案的设计和策划通常需要考虑以下几个方面：

#### 7.3.1.1 参与实验室

- a) 把组织比对的实验室叫做主导实验室，被邀请参与比对的实验室叫做参与实验室；
- b) 如果可能，应确定一个实验室作为参照实验室，假定参照实验室的检测结果准确可靠。将其其他参与实验室的检测结果与参照实验室的结果进行比较，得出参与实验室的检测结果是否准确可靠的结论。

#### 7.3.1.2 检测方法

- a) 所有参与比对的实验室应尽可能使用同一方法
- b) 选用的方法应具有可接受的重复性，不会因为环境的变化和检测人员不同、使用的试剂和消耗性材料不同而产生不可接受的差异；
- c) 所有参与比对的实验室应使用方法规定的统一程序进行检测。主导实验室应在比对开始之前进行必要的试验，也可委托有能力的实验室进行这项试验。例如实施耐用性试验，以获得应进行严格控制

的实验条件。必要时，应征求参与实验室的意见，尽可能在比对开始之前详细了解需要严格控制的检测条件。

### 7.3.1.3 样品

a) 用于比对的样品应具有足够的均匀性和稳定性。

注：必要时，可参照CNAS-GL003-2018《能力验证样品均匀性和稳定性评价指南》的方法实施样品的均匀性和稳定性评价。

b) 如果有必要并可能的情况下，样品的前处理可统一由一个实验室完成。

c) 如果确定参照实验室不可行，应尽可能使用标准样品，如实物标准物质或具有准确的接受参考值的样品；

d) 主导实验室应正确分割样品，适当的包装及适当的样品发送运输方式，确保样品到达参与实验室不会损坏、变质影响检测。

### 7.3.1.4 作业指导书或比对说明

a) 主实验室应编制详细的作业指导书或比对说明，确保参与实验室能按统一的要求和程序实施检测

b) 作业指导书或比对说明至少以下详细内容

1) 收到样品后，实施检查确认的内容，填写样品确认书的要求，样品保管说明；

2) 至少应在收到样品后多少天内完成检测及结果报送的要求；

3) 独立重复检测的样品个数和重复性/再现性条件规定；

4) 样品处置程序及注意事项；

5) 详细的检测程序及重要的注意、控制事项f) 数据采集、计算转换、数据修约及有效位数要求，报送结果数量要求；

6) 报送记录的内容要求；

7) 参与实验室在结果报送文件中，有关特殊情况说明的要求；

8) 联系方式、联系人等；

9) 其他必要的说明（如果有）。

## 7.3.2 比对结果的计算与评价

### 7.3.2.1 $E_n$ 值法

#### 1) 比对结果的计算

实验室间比对结果的计算， $E_n$ 值是最为常用，也是被国际认同的一个典型的用于比对的统计量，它通过将参加实验室与参考实验室的测量结果进行比较并考虑他们的测量不确定度，来评定其校准能力，所表述的是一个标准化的误差，即

$$E_n = \frac{x-X}{\sqrt{U_{lab}^2+U_{ref}^2}} \dots\dots\dots (24)$$

式中：

$x$ ——参加实验室的测量结果；

$X$ ——参考实验室的参考值（指定值）；

$U_{lab}$ ——参加实验室所报告的测量结果的扩展不确定度；

$U_{ref}$ ——参考实验室所报告的参考值（指定值）的扩展不确定度。

$U_{lab}$ 和 $U_{ref}$ 两者的置信概率应相同（一般为95%）。

#### 2) 比对结果的评价

- a) 当 $|E_n| \leq 1$ , 比对结果满意, 通过。  
 b) 当 $|E_n| > 1$ , 比对结果不满意, 不通过。

对出现 $|E_n| > 1$ 的实验室, 必须仔细查找原因, 采取纠正措施。必要时, 可重新进行测量, 否则, 不能参与最后的统计比对。也就是说, 该实验室的该项检测能力实验室间比对不能通过, 即不具备该项检测能力。对出现 $0.7 \leq |E_n| \leq 1$ 的实验室, 建议适时进行系统风险分析, 并采取有效的控制措施

### 7.3.2.2 t 检验法

#### 1) 比对结果的计算

实验室间比对结果的计算,  $E_n$ 值是最为常用, 也是被国际认同的一个典型的用于比对的统计量, 它通过将参加实验室与参考实验室的测量结果进行比较并考虑他们的测量不确定度, 来评定其校准能力, 所表述的是一个标准化的误差, 即

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{S_{\text{合}}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \dots\dots\dots (25)$$

$$S_{\text{合}} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \dots\dots\dots (26)$$

式中:

$\bar{x}_1$ ——第1个实验室测定值的平均值;

$\bar{x}_2$ ——第2个实验室测定值的平均值;

$n_1$ ——第1个实验室的测定次数;

$n_2$ ——第2个实验室的测定次数;

$S_{\text{合}}$ ——合并标准差;

$S_1$ ——第1个实验室测定值的标准差;

$S_2$ ——第2个实验室测定值的标准差;

#### 2) 比对结果的评价

a) 当 $t \geq t_{\text{表}}$ , 表示二组平均值之间有显著性差异。

b) 当 $t < t_{\text{表}}$ , 表示二组平均值之间无显著性差异。

其中 $t_{\text{表}}$ 是根据之前设定的显著性水平 $\alpha$ , 查自由度 $n_1 + n_2 - 2$ 的“t检验临界值表”得到

## 附录 A (资料性) 数据处理

### A.1 数据的统计处理

几乎所有的离群值的统计检查都是基于这个一个假设，即数据其内在的分布为正态分布。离群值就是与样品中其他数据有很大偏离的数据，它出现在数据组的两个极端。离群值有已经有或发现有事实原因的离群值和不能发现事实原因的离群值。如操作者发现了操作方法偏离了规定的程序，或在计算或记录数据时产生的误差，这类数据就是有事实原因的离群值，必须剔除，而不需进行离群值的统计检验。发现不了事实原因的异常值或离群值，需要进行离群值的统计检验加以确认。

离群值统计检验适用于指定材料某种特性的重复测试结果或随机样本的多次测试结果的检验。

#### A.1.1 离群值的处理

##### A.1.1.1 单个离群值的检验

##### A.2.1.1.1 格拉布斯 (Grubbs) 检验

###### a) 建立 Grubbs 统计量

假如单一样品进行n次重复测试，得n个测试结果数据，按其按数值大小升序排列，设为 $x_1, x_2, x_3, \dots, x_n$ ，怀疑最大一个即 $x_n$ 有问题，建立Grubbs统计量 $G_n$ ：

$$G_n = (x_n - \bar{x})/S \quad \dots\dots\dots (A1)$$

式中：

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad \dots\dots\dots (A2)$$

反之，如怀疑最小值即 $x_1$ 有问题，Grubbs统计量变为 $G_1$ ：

$$G_1 = (\bar{x} - x_1)/S \quad \dots\dots\dots (A3)$$

###### b) 判定准则

- 1) 假如该检验统计量小于等于其 5%临界值，则接受被检验项目为合格；
- 2) 假如该检验统计量大于其 5%但小于或等于其 1%临界值，则称被检验项目歧离值；
- 3) 假如该检验统计量大于其 1%临界值，则被检验项目称为统计离群值。

临界值见下表。

###### c) 适用范围

- 1) 不同数据组的平均值的离群值检验

如能力验证的组织者组织了对同一个样品进行测试的能力验证，共由n个实验室参加，每个实验室按要求产生一组数据，其平均值按升序排列分别为 $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n$ ，组织者如怀疑某个最大或最小平均值是离群值，可用上述公式进行检验，但用 $\bar{x}_n$ 或 $\bar{x}_1$ 代替 $x_n$ 或 $x_1$ ，用 $\bar{\bar{x}}$ 代替 $\bar{x}$ ，其中 $\bar{\bar{x}}$ 所平均值的平均值：

$$\bar{\bar{x}} = \frac{\sum_{i=1}^n \bar{x}_i}{n} \quad \dots\dots\dots (A4)$$

- 2) 同一组数据中的单个离群值的检验。

###### d) 单个离群值检验的递归应用

对一组数据的一个极端数据按上述方法进行离群值检验并被判为离群值及被剔除后，如对余下数据还有怀疑，则就对该数据组的另一极端数据或同一端的其他数据进行相同的检验，如被判为离群值并被

剔除后，对余下数据还有怀疑，则继续进行上述从一个极端到另一个极端或同一端数据的检验过程，一直没有发现离群值。

递归应用也适用于上述第3条的两种情况。

#### A. 1. 1. 2 两个离群值的同时检验

对同一数据组的同一端两个数据或两端的各一个数据有怀疑时，可分别使用同一端两个离群值的同时检验和两端各1个离群值的同时检验。

#### A. 1. 1. 3 同一端两个离群值的同时检验

##### A. 1. 1. 3. 1 格拉布斯 (Grubbs) 检验

两个最大的数据有怀疑时，建立Grubbs统计量G：

$$G = S_{n-1,n}^2 / S_o^2 \quad \dots\dots\dots (A5)$$

式中：

$$S_o^2 = \sum_{i=1}^n (x_i - \bar{x})^2 \quad \dots\dots\dots (A6)$$

$$S_{n-1,n}^2 = \sum_{i=1}^{n-2} (x_i - \bar{x}_{n-1,n})^2 \quad \dots\dots\dots (A7)$$

$$\bar{x}_{n-1,n} = \frac{\sum_{i=1}^{n-2} x_i}{n-2} \quad \dots\dots\dots (A8)$$

以同样的方式建立两个最小数据的Grubbs统计量G：

$$G = S_{1,2}^2 / S_o^2 \quad \dots\dots\dots (A9)$$

式中：

$$S_{1,2}^2 = \sum_{i=3}^n (x_i - \bar{x}_{1,2})^2 \quad \dots\dots\dots (A10)$$

$$\bar{x}_{1,2} = \frac{\sum_{i=3}^n x_i}{n-2} \quad \dots\dots\dots (A11)$$

将统计量G与下表的临界值进行比较以决定是否同时存在离群值。

Table 5 — Critical values for Grubbs' test

<i>p</i>	One largest or one smallest		Two largest or two smallest	
	Upper 1 %	Upper 5 %	Lower 1 %	Lower 5 %
3	1,155	1,155	—	—
4	1,496	1,481	0,000 0	0,000 2
5	1,764	1,715	0,001 8	0,009 0
6	1,973	1,887	0,011 6	0,034 9
7	2,139	2,020	0,030 8	0,070 8
8	2,274	2,126	0,056 3	0,110 1
9	2,387	2,215	0,085 1	0,149 2
10	2,482	2,290	0,115 0	0,186 4
11	2,564	2,355	0,144 8	0,221 3
12	2,636	2,412	0,173 8	0,253 7
13	2,699	2,462	0,201 6	0,283 6
14	2,755	2,507	0,228 0	0,311 2
15	2,806	2,549	0,253 0	0,336 7
16	2,852	2,585	0,276 7	0,360 3
17	2,894	2,620	0,299 0	0,382 2
18	2,932	2,651	0,320 0	0,402 5
19	2,968	2,681	0,339 8	0,421 4
20	3,001	2,709	0,358 5	0,439 1
21	3,031	2,733	0,376 1	0,455 6
22	3,060	2,758	0,392 7	0,471 1
23	3,087	2,781	0,408 5	0,485 7
24	3,112	2,802	0,423 4	0,499 4
25	3,135	2,822	0,437 6	0,512 3
26	3,157	2,841	0,451 0	0,524 5
27	3,178	2,859	0,463 8	0,536 0
28	3,199	2,876	0,475 9	0,547 0
29	3,218	2,893	0,487 5	0,557 4
30	3,236	2,908	0,498 5	0,567 2
31	3,253	2,924	0,509 1	0,576 6
32	3,270	2,938	0,519 2	0,585 6
33	3,286	2,952	0,528 8	0,594 1
34	3,301	2,965	0,538 1	0,602 3
35	3,316	2,979	0,546 9	0,610 1
36	3,330	2,991	0,555 4	0,617 5
37	3,343	3,003	0,563 6	0,624 7
38	3,356	3,014	0,571 4	0,631 6
39	3,369	3,025	0,578 9	0,638 2
40	3,381	3,036	0,586 2	0,644 5

Reproduced, with the permission of the American Statistical Association, from reference [4] in annex C.

*p* = number of laboratories at a given level

## A. 1. 1. 4 两端各 1 个离群值的同时检验

## A. 1. 1. 4. 1 David, Hartley 和 Pearson 检验

当对两端各 1 个数据（即最小和最大）有怀疑时，按照 David, Hartley 和 Pearson 的样本标准偏差检验建立如下公式：

$$w/S = (x_n - x_1)/S \dots\dots\dots (A12)$$

将上述极差对标准偏差的比例 ( $w/s$ ) 与下表临界值进行比较, 以决定是否同时存在2个离群值。

**TABLE 3 Critical Values<sup>A</sup> (One-Sided Test) for  $w/s$  (Ratio of Range to Sample Standard Deviation)**

Number of Observations, $n$	10 Percent Significance Level	5 Percent Significance Level	1 Percent Significance Level
3	1.9973	1.9993	2.0000
4	2.409	2.429	2.445
5	2.712	2.755	2.803
6	2.949	3.012	3.095
7	3.143	3.222	3.338
8	3.308	3.399	3.543
9	3.449	3.552	3.720
10	3.574	3.685	3.875
11	3.684	3.803	4.011
12	3.782	3.909	4.133
13	3.871	4.005	4.244
14	3.952	4.092	4.344
15	4.025	4.171	4.435
16	4.093	4.244	4.519
17	4.156	4.311	4.597
18	4.214	4.374	4.669
19	4.269	4.433	4.736
20	4.320	4.487	4.799
21	4.368	4.539	4.858
22	4.413	4.587	4.913
23	4.456	4.633	4.965
24	4.497	4.676	5.015
25	4.535	4.717	5.061
26	4.572	4.756	5.106
27	4.607	4.793	5.148
28	4.641	4.829	5.188
29	4.673	4.863	5.226
30	4.704	4.895	5.263
35	4.841	5.040	5.426
40	4.957	5.162	5.561
45	5.057	5.265	5.674
50	5.144	5.356	5.773

#### A. 1. 1. 5 稳健的统计方法

稳健的统计方法 (**robust statistical method**) 是对给定概率模型基本假设条件微小偏离不敏感的统计方法, 即不受离群值影响的统计方法, 在利用稳健的统计方法进行均值和标准偏差计算时不需要剔除离群值, 因此, 当怀疑数据中有离群值又不好处理时, 最好使用稳健的数据分析方法。

ISO 5725-5 给出了运算法则 A (**Algorithm A**) 和运算法则 S (**Algorithm S**) 的稳健统计方法, 其中运算法则 A 用来计算均值、标准偏差, 运算法则 S 用来计算合并标准偏差。

#### 1、运算法则 A (**Algorithm A**)

##### (1) 数据组

假如一组 $n$ 个测试结果数据,按其按数值大小升序排列,设为 $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ ,  
设其中位值为 $med(x)$ 、按稳健统计方法计算得到的均值为 $x^*$ ,标准偏差为 $s^*$

## (2) 计算中位值 $med(x)$

当 $n$ 为奇数时,  $med(x) = \frac{x_{n+1}}{2}$  ..... (A13);

当 $n$ 为偶数时,  $med(x) = \frac{\frac{x_n+x_{(1+\frac{n}{2})}}{2}}{2}$  ..... (A14);

## (3) 计算 $x^*$ 、 $s^*$ 的初值:

$$x^* = med(x) \quad \dots\dots\dots (A15)$$

设 $d_i = |x_i - x^*|$ ,  $med(d)$ 为 $d_1, d_2, d_3, \dots, d_i, \dots, d_n$ 的中位值;

$$\text{则, } s^* = 1.483 \times med(d) \quad \dots\dots\dots (A16)$$

式中:  $i = 1, 2, 3, \dots, n$

## (4) 更新 $x^*$ 、 $s^*$

计算 $\varphi$

$$\varphi = 1.5s^* \quad \dots\dots\dots (xx)$$

对 $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ , 中的每一个值 $x_i$ , 计算下列值:

$$x'_i = \begin{cases} x^* - \varphi & \text{假如 } x_i < x^* - \varphi \\ x^* + \varphi & \text{假如 } x_i < x^* + \varphi \\ \text{其他情况, } x_i & \end{cases} \quad \dots\dots\dots (A17)$$

按照下式计算新的 $x^*$ 、 $s^*$

$$x^* = \sum_{i=1}^n x'_i / n \quad \dots\dots\dots (A18)$$

$$s^* = 1.134 \sqrt{\sum_{i=1}^n (x'_i - x^*)^2 / (n - 1)} \quad \dots\dots\dots (A19)$$

将新计算获得的 $x^*$ 、 $s^*$ 按(4)步再来一次,再次获得新的 $x^*$ 、 $s^*$ ,此迭代过程不断持续,  
直到 $x^*$ 、 $s^*$ 的前三位有效数字保持不变后,迭代过程结束。

## 2、运算法则 S (Algorithm S)

运算法则 S (Algorithm S) 是用来计算合并标准偏差或极差的。

### (1) 数据组

假如一组 $n$ 个标准偏差或极差数据,按其按数值大小升序排列,设为 $w_1, w_2, w_3, \dots, w_i, \dots, w_n$ , 设其中位值为 $med(w)$ 、按稳健统计方法计算得到的合并标准偏差或极差值为 $w^*$ 。



(2) 限值系数 $\eta$  (Limit factor) 和调整系统 $\xi$  (Adjustment factor)

通过使用该数据组的每个组成数据的自由度 $\nu$ 来查表(表 F.1) 获得稳健统计分析所使用的限值系数 $\eta$  (Limit factor) 和调整系统 $\xi$  (Adjustment factor), 本文假设每个组成数据的获得是基于相同数量 $p$ 的重复测量。当组成数据为标准偏差时, 自由度为 $\nu = p - 1$ ; 当组成数据为极差时, 自由度为 $\nu = 1$ 。

表 F.1 稳健统计分析所用的系数

自由度 $\nu$	限值系数 $\eta$	调整系数 $\xi$
1	1.645	1.097
2	1.517	1.054
3	1.444	1.039
4	1.395	1.032
5	1.359	1.027
6	1.332	1.024
7	1.310	1.021
8	1.292	1.019
9	1.277	1.018
10	1.264	1.017

(3) 计算 $w^*$ 的初值

$$w^* = \text{med}(w) \quad \dots\dots\dots (A20)$$

(4) 更新 $w^*$ 的值

$$\psi = \eta \times w^* \quad \dots\dots\dots (A21)$$

对 $w_1, w_2, w_3, \dots, w_i, \dots, w_n$ 中的每一个值 $w_i$ , 计算下列值:

$$w'_i = \begin{cases} \psi & \text{假如 } w_i > \psi \\ w_i & \text{其他情况} \end{cases} \quad \dots\dots\dots (A22)$$

计算新的 $w^*$ :

$$w^* = \xi \sqrt{\sum_{i=1}^n (w'_i)^2 / n} \quad \dots\dots\dots (A23)$$

将新计算获得的 $w^*$ 按(4)步再来一次, 再次获得新的 $w^*$ , 此迭代过程不断持续, 直到 $w^*$

的前三位有效数字保持不变后，迭代过程结束。

如 $w^*$ 为极差的稳健计算结果时，按下式转换成标准偏差SD

$$SD = w^*/\sqrt{2} \quad \dots\dots\dots (A24)$$

## A. 1. 2 正态总体均值的假设检验（参数检验）

### A. 1. 2. 1 单个正态总体均值 $\mu$ 的检验

设显著性水平为 $\alpha$ ，正态总体均值 $\mu$ 为已知数 $\mu_0$ 。

1. U检验：总体 $X \sim N(\mu, \sigma^2)$ ， $\sigma^2$ 已知的情形。

设 $X_1, X_2, \dots, X_n$ 是取自正态总体 $X \sim N(\mu, \sigma^2)$ 的样本，其统计量U从标准正态分布，即：

$$U \sim N(0, 1)$$

其中：

$$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

使用时，将计算所得的U的绝对值与正态分布表中在置信水平为95%时所查得的值 $z_{\frac{\alpha}{2}}$ （也叫第95个百分位数）进行比较，如大于所查的值，即：

$$\frac{|\bar{X} - \mu|}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}}$$

则：

$$|\bar{X} - \mu| > z_{\frac{\alpha}{2}} \cdot \sigma/\sqrt{n}$$

实验室的值与 $\mu$ 值有显著差异。否则，两者没有显著差异。

在具体使用时，还要考虑下列两种情况：

① 当 $\bar{X}$ 与标准物质的 $\mu$ 值进行比较时，如标准物质本身的测量不确定度大于 $0.3\sigma/\sqrt{n}$ 时，需要考虑标准物质的测量不确定度 $u_{ref}$ ，上述式子变成：

$$U = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n} + (u_{ref})^2}}$$

② 在使用标准方法测试时，要使用实验室的再现性方差 $\sigma_R^2$ 进行结果评定，因此，除了重复性方差 $\sigma_r^2$ 外，还需要考虑实验室间偏差的方差 $\sigma_L^2$ 。为此，上述U统计量公式不能直接使用。

$$\sigma_R^2 = \sigma_L^2 + \sigma_r^2$$

由于是重复测试， $\sigma_r^2$ 要用 $\sigma_r^2/n$ 代替，因此 $\sigma_R^2$ 公式变成：

$$\sigma_R^2 = \sigma_L^2 + \sigma_r^2/n$$

所以， $\bar{X}$ 满足方差为 $\sigma_R^2$ 的正态分布：

$$\bar{X} \sim N(\mu, \sigma_R^2)$$

将上述转换成标准正态分布：

$$\frac{\bar{X} - \mu}{\sigma_R} \sim N(0,1)$$

在这种情况下，U 值变为：

$$U = \frac{\bar{X} - \mu}{\sigma_R}$$

同理，将 U 绝对值与置信水平为 95% 的标准正态分布的第 95 个百分位数进行比较。经查第 95 个百分位数为 1.96，约等于 2。因此，

$$\frac{|\bar{X} - \mu|}{\sigma_R} < 2$$

$$|\bar{X} - \mu| < 2\sigma_R = 2\sqrt{\sigma_L^2 + \sigma_r^2/n} = 2\sqrt{\sigma_L^2 + \sigma_r^2 - \sigma_r^2(1 - \frac{1}{n})}$$

$$|\bar{X} - \mu| < 2\sqrt{\sigma_R^2 - \sigma_r^2(1 - \frac{1}{n})} = \frac{1}{\sqrt{2}}\sqrt{8\sigma_R^2 - 8\sigma_r^2(1 - \frac{1}{n})}$$

$$|\bar{X} - \mu| < \frac{1}{\sqrt{2}}\sqrt{(2.8\sigma_R)^2 - (2.8\sigma_r)^2(\frac{n-1}{n})}$$

$$|\bar{X} - \mu| < \frac{1}{\sqrt{2}}\sqrt{(R)^2 - (r)^2(\frac{n-1}{n})}$$

其中：R 为再现性限，r 为重复限。

设：

$$CD = \frac{1}{\sqrt{2}}\sqrt{(R)^2 - (r)^2(\frac{n-1}{n})}$$

把 CD 叫做差值在置信水平为 95% 的临界值。也即实验室测试的平均值与标准值相比不能大于 CD 值，如大于 CD，则它们之间有显著差异，否则就没有显著差异。

如标准物质的测量不确定度大于  $\sigma_R$  时，需要考虑标准物质的测量不确定度  $u_{ref}$ ，上述公式 (XXX) 变为：

$$\frac{|\bar{X} - \mu|}{\sqrt{\sigma_R^2 + (u_{ref})^2}} < 2$$

$|\bar{X} - \mu|$  和 CD 分别变为：

$$|\bar{X} - \mu| < \frac{1}{\sqrt{2}}\sqrt{(R)^2 - (r)^2(\frac{n-1}{n}) + (2.8u_{ref})^2}$$

$$CD = \frac{1}{\sqrt{2}}\sqrt{(R)^2 - (r)^2(\frac{n-1}{n}) + (2.8u_{ref})^2}$$

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

例1 某实验室委托某单位加工浓度为 5.0 (单位略) 的一批能力验证样品，为了验证这批样品的质量情况，实验室抽取了 10 个样进行检验，结果如下：

5.1, 5.0, 4.9, 4.8, 5.1, 4.9, 4.8, 4.9, 5.0, 4.9,

按以往的经验,能力验证样品浓度 $X$ 的分布服从正态分布 $X \sim N(\mu, \sigma^2)$ , 其中 $\sigma = 0.05$ , 如何判断这批样品的浓度是否符合要求?

计算过程如下:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = 4.94$$

$$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{4.94 - 5.0}{0.05/\sqrt{10}} = -3.8$$

经查正态分布在置信水平为95%时的数值 $z_{0.025} = 1.96$ , 绝对值3.8大于1.96, 因此:

$$|\bar{X} - \mu| > z_{\frac{\alpha}{2}} \cdot \sigma/\sqrt{n} = 1.96 \cdot \frac{0.05}{\sqrt{10}} = 0.03$$

而 $|\bar{X} - \mu|$ 值为0.06, 远大于0.03, 因此, 浓度均值与委托加工的浓度值有显著差异, 所加工的浓度不符合要求。

(2) t检验: 总体 $X \sim N(\mu, \sigma^2)$ ,  $\sigma^2$ 未知的情形。

就是用服从t分布的统计量来确定拒绝域, 从而对正态总体均值进行检验的方法。

由于 $\sigma^2$ 未知, 所以不能用上述U统计量来计算, 但 $S^2$ 是可计算得到的, 按照定理, 由于统计量

$\frac{\bar{X} - \mu}{S/\sqrt{n}}$ 服从自由度 $n-1$ 的t-分布, 所以可用t检验来进行。

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$

在指定显著性水平 $\alpha$ 下, 如 $|t| \geq t_{\frac{\alpha}{2}}(n-1)$ , 则 $\bar{X}$ 与 $\mu_0$ 有显著差异。

**例2** 用某方法进行测试时, 某实验室利用Pb值为100 mg/kg的CRM (有证标准物质) 物质进行回收率实验, 测试结果如下 (单位: mg/kg):

98, 99, 97, 96, 98, 96, 97, 97

通过对回收率结果的分析, 请问从是否需要日常测试的结果进行修正? (假设测试结果分布呈正态分布,  $\alpha=0.05$ )

说明: 本题从8次测量结果来判断所有的测量值 (作为一个总体) 的均值 $\mu$ 是否等于100 mg/kg。属于双侧检验。由于没有给出 $\sigma$ 值, 因此可用t检验进行。

解:

统计量:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$

给出显著性水平 $\alpha$ , 如0.05, 查得t的临界值。

按照计算结果, 如t绝对值大于临界值 $t_{0.025}$ , 则测试结果均值与CRM的值之间有显著差异。计算过程如下:

$\bar{x}=97.3$ ,  $S=1.0$ , 属于双侧检验, 查t分布表(自由度为7)得 $t_{0.025}=2.36$ , 所以

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{97.3 - 100}{1.0/\sqrt{8}} = -7.6$$

$$|t| = 7.6 > 2.36$$

所以, 测量结果均值与CRM的值之间有显著差异, 需要对测试结果进行修正。

**例3** 某工厂生产涂料, 含铅量为50mg/kg, 为了将含铅量降下来, 工厂改进了生产工艺, 为了验证是否降低了铅含量, 实验室抽取了10样进行检验, 结果如下:

48, 49, 48, 49, 47, 48, 49, 49, 49, 48

问新工艺是否真的降低涂料中的铅含量? (假设含量分布呈正态分布,  $\alpha=0.05$ )

说明: 本题从10个检验结果来判断该批涂料(作为一个总体)的均值 $\mu$ 是否小于50 mg/kg。属于左侧检验。由于没有给出 $\sigma$ 值, 因此可用t检验进行。

解:

统计量:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$

给出显著性水平 $\alpha$ , 如0.05, 查得t的临界值。

按照计算结果, 如结果落t值小于临界值 $t_{0.05}$ 的负值(对于t分布, 由于是对称分布, 左侧的临界值等于右侧临界值的负值。查表一般只能得到右边临界值), 则测试结果与50mg/kg有显著差异, 也就是铅含量已降了下来。

计算过程如下:

$\bar{x}=48.4$ ,  $S=0.7$ , 属于单侧(左侧)检验, 查t分布表(自由度为9)得 $t_{0.05}=1.8331$ ,

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{48.4 - 50}{0.7/\sqrt{10}} = -7.2$$

由于计算所得的t值小于查表所得的值的负值, 则测试结果与50mg/kg有显著差异, 铅含量确实有所降低。

#### A. 1. 2. 2 两个正态总体均值是否相等的检验

上面是讲到对单个正态总体 $\mu$ 的判断问题, 有时会遇到要判断两个正态总体的 $\mu$ 之间是否相等的问题, 如实验室常用质量监控样对实验室的质量进行控制, 当上一批质控样快用完时, 往往要制备或新购进一批质控样, 但新购或制备的质控制样是否与上一批一样, 这时就要进行判断。

设 $X_1, X_2, \dots, X_m$ 为来自总体 $X \sim N(\mu_1, \sigma_1^2)$ 的样本, 而 $Y_1, Y_2, \dots, Y_n$ 为来自总体 $Y \sim N(\mu_2, \sigma_2^2)$ 的样本,  $X$ 与 $Y$ 相互独立。

(1) U检验: 方差 $\sigma_1^2$ 和 $\sigma_2^2$ 已知的情况

假设 $H_0: \mu_1 = \mu_2$ ; ;

$$\text{构建统计量: } U = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$$

由于假设 $\mu_1 = \mu_2$ ; , 所以

$$U = \frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$$

设显著水平为 $\alpha$ , 如:

$$\bar{X} - \bar{Y} \leq -U_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} \text{ 或 } \bar{X} - \bar{Y} \geq U_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}, \text{ 则 } \bar{X} \text{ 与 } \bar{Y} \text{ 之间有显著差异。}$$

(2) **t检验:** 方差 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ , 但 $\sigma^2$ 未知的情况

假设 $H_0: \mu_1 = \mu_2$ :

根据定理, 统计量:

$$t = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{1/m + 1/n}} \sim t(m+n-2)$$

$$\text{其中: } S_w^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}$$

设显著水平为 $\alpha$ , 如:

$$\bar{X} - \bar{Y} \geq t_{\frac{\alpha}{2}}(m+n-2) \cdot S_w \sqrt{1/m + 1/n}$$

或

$$\bar{X} - \bar{Y} \leq -t_{\frac{\alpha}{2}}(m+n-2) \cdot S_w \sqrt{1/m + 1/n}, \text{ 则 } \bar{X} \text{ 与 } \bar{Y} \text{ 之间有显著差异。}$$

**例4** 某实验室分别用美国CPSC制定的方法和中国国家标准方法分别测定同一批塑料中的增塑剂DEHP, 测量结果如下 (单位为mg/kg):

CPSC方法: 135, 138, 140, 145, 134, 136, 132, 147;

国家标准方法: 150, 145, 155, 156, 143, 158, 155, 150。

请问两个平均值间是否有显著差异? (假设测量结果分布呈正态分布,  $\alpha=0.05$ )

说明: CPSC方法所得的测试结果与国家标准方法所得的测试结果可看作是不同的总体, 因此可用上

面的公式来检验。由于没有给出 $\sigma^2$ 值, 因此可用t检验进行。

解: 在进行均值检验前, 要先检验两者的方差是否相等, 相等后才能应用上述t检验。

经F检验可知 (见下节), 两者的方差相等, 故可以建立以下的假设检验:

统计量:

$$t = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{1/m + 1/n}} \sim t_{m+n-2}$$

给出显著性水平 $\alpha$ ，如0.05，查得t的临界值。

按照计算结果，如结果t的绝对值大于临界值 $t_{0.025}$ （属于双侧检验），则 $\bar{X}$ 与 $\bar{Y}$ 之间有显著差异。计算过程如下：

$\bar{x}=138.4$ ， $\bar{y}=151.5$ ， $S_1=5.3$ ， $S_2=5.4$ ，按下式计算 $S_w$ ：

$$S_w^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2} = \frac{(8-1) \times 5.3^2 + (8-1) \times 5.4^2}{8+8-2} = 28.6$$

$$t = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{1/m + 1/n}} = \frac{138.4 - 151.5}{5.3 * \sqrt{1/8 + 1/8}} = -4.9$$

查t分布表（自由度为14）得 $t_{0.025}=2.1$ ，由于计算所得的t值的绝对值大于查表所得的值，故两者的平均值有显著差异。

在本例中，如已给出方差值，假设 $\sigma_1^2=25$ 和 $\sigma_2^2=30$ ，则可按U检验进行。

#### A. 1. 2. 3 两个正态总体方差是否相等的假设检验——F 检验

在例4中，对按两个标准方法所进行的测试结果的均值是否一致进行了检验，在进行该项检验前，要求对两个总体的方差是否相等进行检验，下面介绍有关的检验方法。

设 $X_1, X_2, \dots, X_m$ 为来自总体 $X \sim N(\mu_1, \sigma_1^2)$ 的样本，而 $Y_1, Y_2, \dots, Y_n$ 为来自总体 $Y \sim N(\mu_2, \sigma_2^2)$ 的样本， $X$ 与 $Y$ 相互独立。

根据定理，统计量：

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{m-1, n-1}$$

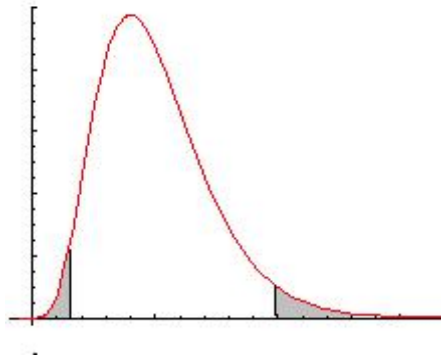
由于假设 $\sigma_1^2 = \sigma_2^2$ ，所以：

$$F = \frac{S_1^2}{S_2^2} \sim F_{m-1, n-1}$$

设显著水平为 $\alpha$ ，如：

$$\frac{S_1^2}{S_2^2} \leq F_{1-\frac{\alpha}{2}}(m-1, n-1) \text{ 或 } \frac{S_1^2}{S_2^2} \geq F_{\frac{\alpha}{2}}(m-1, n-1), \text{ , 则两个方差之间有显著差异。}$$

见图2-19。



$$F_{1-\frac{\alpha}{2}}(m-1, n-1) \quad F_{\frac{\alpha}{2}}(m-1, n-1)$$

图2-19 F分布的双侧检验

在F分布表中，一般只列出  $F_{\frac{\alpha}{2}}(m-1, n-1)$ ，而不列出  $F_{1-\frac{\alpha}{2}}(m-1, n-1)$  值，但后者可通过下列转换得到：

$$F_{\frac{\alpha}{2}}(n-1, m-1) = \frac{1}{F_{1-\frac{\alpha}{2}}(m-1, n-1)} \quad (2-7)$$

例如： $\alpha=0.05$ ， $m=10$ ， $n=8$ 时，如属于双侧检验，查F分布表得：

$$F_{0.025}(9, 7) = 4.82，\text{ 而 } F_{0.975}(9, 7) \text{ 是查不到的，但可查得 } F_{0.025}(7, 9) = 4.20，$$

通过式 (2-7) 的可得  $F_{0.975}(9, 7)$ ：

$$F_{0.975}(9, 7) = \frac{1}{F_{0.025}(7, 9)} = \frac{1}{4.20} = 0.24$$

**例5** 用F检验来说明上述例4的方差是否相等？

解：

统计量：

$$F = \frac{S_1^2}{S_2^2} \sim F_{m-1, n-1}$$

(iii) 给出显著性水平  $\alpha$ ，如0.05，属于双侧检验，由于F分布是不对称的，要先查出  $F_{0.025}(7, 7)$  的值，然后再转换成  $F_{0.975}(7, 7)$ ；

当计算结果显示：

$$\frac{S_1^2}{S_2^2} \geq F_{\frac{\alpha}{2}}(m-1, n-1) \text{ 或 } \frac{S_1^2}{S_2^2} \leq F_{1-\frac{\alpha}{2}}(m-1, n-1)$$

则两个方差之间有显著差异。

计算过程如下：

通过计算得  $S_1^2 = 28.1$ ， $S_2^2 = 29.2$ ，F值等于0.96，查表得  $F_{0.025}(7, 7) = 4.99$ ，按式 (2-7) 得  $F_{0.975}(7, 7)$ ：



$$F_{0.975}(7,7) = \frac{1}{F_{0.025}(7,7)} = \frac{1}{4.99} = 0.20$$

由于F值小于4.99并且大于0.20，即两个方差间没有显著差异。

### A.1.3 多台设仪器比较

在重复性条件下，但让其中的仪器设备发生变化，也即在中间精密度条件下，对 $p$ 台仪器设备进行测试，每台仪器设备的测试次数均为 $n$  ( $n \geq 6$ )，得到各台仪器设备的平均值为 $\bar{y}_1, \bar{y}_2, \bar{y}_3, \dots, \bar{y}_p$ 。

已知实验室设备测试的测量不确定度为 $u$ ，假设测试结果的总平均值为 $\bar{y}$ ，则：

$$\bar{y} = \frac{\sum_{i=1}^p \bar{y}_i}{p}$$

由于 $\bar{y}_i$ 的标准不确定度 $u_{\bar{y}_i} = \frac{u}{\sqrt{n}}$ ，所以 $\bar{y}$ 的标准不确定度 $u_{\bar{y}}$ 等于：

$$u_{\bar{y}} = \sqrt{\frac{1}{p^2}(p \cdot u_{\bar{y}_i}^2)} = \sqrt{\frac{1}{p}(u_{\bar{y}_i}^2)} = \frac{u_{\bar{y}_i}}{\sqrt{p}} = \frac{u}{\sqrt{pn}}$$

因此，设 $(\bar{y}_i - \bar{y})$ 的不确定度为 $u_y$ ，则：

$$u_y = \sqrt{u_{\bar{y}_i}^2 + u_{\bar{y}}^2} = u \sqrt{\frac{1}{n}(1 + \frac{1}{p})}$$

但该公式是不准确的，因为， $\bar{y}_i$ 与 $\bar{y}$ 不是相互独立的，不能采用该公式，也不能采用下述公式：

$$|\bar{y}_i - \bar{y}| < ku \sqrt{\frac{1}{n}(1 + \frac{1}{p})}$$

但如果 $\bar{y}$ 采用历史数据，而不是使用当次数据，则可能避开独立性的问题，因此，上述过程修改如下：

在重复性条件下，但让其中的仪器设备发生变化，也即在中间精密度条件下，对 $p$ 台仪器设备进行测试，每台仪器设备的测试次数均为 $n$  ( $n \geq 6$ )，得到各台仪器设备的平均值为 $\bar{y}_1, \bar{y}_2, \bar{y}_3, \dots, \bar{y}_p$ 。

假设测试结果的总平均值为 $\bar{y}$ ，已知实验室设备测试的测量不确定度为 $u$ ，已知以往总平均值的数据值为 $\bar{y}_0$ ，其标准不确定度为 $u_{\bar{y}_0}$ 。

由于 $\bar{y}_i$ 的标准不确定度 $u_{\bar{y}_i} = \frac{u}{\sqrt{n}}$ ，设 $(\bar{y}_i - \bar{y}_0)$ 的不确定度为 $u_y$ ，则：

$$u_y = \sqrt{u_{\bar{y}_i}^2 + u_{\bar{y}_0}^2} = \sqrt{\frac{1}{n}u^2 + u_{\bar{y}_0}^2}$$

所以，

$$|\bar{y}_i - \bar{y}_0| < k \sqrt{\frac{1}{n}u^2 + u_{\bar{y}_0}^2}$$

取 $k = 2$ ，上式变成：

$$|\bar{y}_i - \bar{y}_0| < 2 \sqrt{\frac{1}{n} u^2 + u_{\bar{y}_0}^2}$$

一般很少一个仪器的平均值与总的平均值来比较（如上所述），如要比较，可使用格拉布斯（Grubbs）检验可能更加有效，其将数据组中的最大或最小与当次测试的平均值进行比较，然后将离群值剔除，如无离群值，则数据组中的所有数据均符合。另外一种方式，就是将本次测试的总的平均值与以往的平均值进行比较，即： $|\bar{y} - \bar{y}_0|$ 。

$$\bar{y} = \frac{\sum_{i=1}^p \bar{y}_i}{p}$$

$$u_{\bar{y}} = \sqrt{\frac{1}{p^2} (p \cdot u_{\bar{y}_i}^2)} = \sqrt{\frac{1}{p} (u_{\bar{y}_i}^2)} = \frac{u_{\bar{y}_i}}{\sqrt{p}} = \frac{u}{\sqrt{pn}}$$

设 $(\bar{y} - \bar{y}_0)$ 的标准不确定为 $u_{y'}$

$$u_{y'} = \sqrt{u_{\bar{y}}^2 + u_{\bar{y}_0}^2} = \sqrt{\frac{u^2}{pm} + u_{\bar{y}_0}^2}$$

$$|\bar{y} - \bar{y}_0| < k \sqrt{\frac{u^2}{pm} + u_{\bar{y}_0}^2}$$

取 $k = 2$ , 上式变成:

$$|\bar{y} - \bar{y}_0| < 2 \sqrt{\frac{u^2}{pm} + u_{\bar{y}_0}^2}$$

$|\bar{y} - \bar{y}_0|$ 如符合要求，则这次检测的所有设备均符合要求，如不符合，进行格拉布斯（Grubbs）检验，将不符合的一组数据找出来，并将该组数据排除之外，再把剩下的数据按上式计算 $|\bar{y} - \bar{y}_0|$ ，如符合，则其他设备的测试符合要求，不然再进行格拉布斯（Grubbs）检验，排除数据，再进行 $|\bar{y} - \bar{y}_0|$ ，此过程一直进行直到问题解决。

#### A2.4 卡方检验

设 $X_1, X_2, \dots, X_n$ 是取自正态总体 $X \sim N(\mu, \sigma^2)$ 的样本，则：

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

其中：

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

设统计量 $\chi^2$ 为下式：

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

在指定显著性水平 $\alpha$ 下，当 $\chi^2 < \chi_{n-1}^2(1 - \frac{\alpha}{2})$ 或 $\chi^2 > \chi_{n-1}^2(\frac{\alpha}{2})$ 时， $S^2$ 与 $\sigma^2$ 之间均有显著性差异。

用EXCEL制作控制图

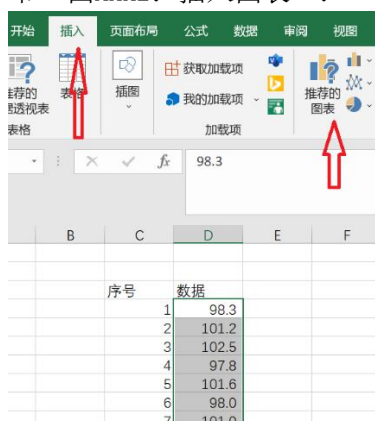
如用实验室内部质量监控样，在15天内测试20次，结果数据见表AXX1。

表AXX1，质控样数据

序号	数据	中位线及控制 限制	
1	98.3	平均值	100.5
2	101.2	标准偏差S	2.0
3	102.5	平均值+2S	104.50
4	97.8	平均值-2S	96.50
5	101.6	平均值+3S	106.50
6	98.0	平均值-3S	94.50
7	101.0		
8	102.8		
9	98.2		
10	99.1		
11	101.5		
12	98.5		
13	104.1		
14	103.1		
15	99.6		
16	101.5		
17	100.1		
18	99.8		
19	98.3		
20	103.6		

按6.5.4“数据的处理”要求，将这些数据在XY坐标上标出来，得到趋势图，再从图上看是否有异排列的数据，再进行离群值检验，接着对这些数据进行正态性检验，包括画正态概率图和用the Anderson-Darling 统计量进行检验，经检查均符合上述要求，然后求出平均值、标准偏差、上下警告限、上下行动限、接着对这些数据制作控制图，步骤如下：

打开EXCEL将上表第一、二栏数据复制到EXCEL表上，将数据一栏选上，不用选序号一栏，点“插入”、再点“推荐的图表”，见“图AXX1 数据”和“图AXX2：插入图表”：



图AXX1 数据

## 附录 B (资料性) 用 EXCEL 制作控制图



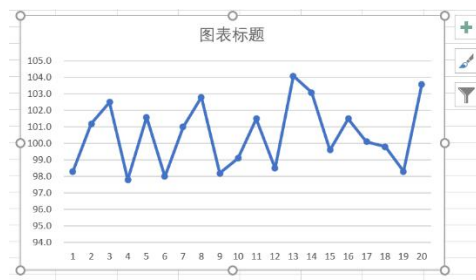
图AXX2: 插入图表

依次点“所有图表”、“折线图”、“带数据标志的折线图”，见“图AXX3: 图表选择”



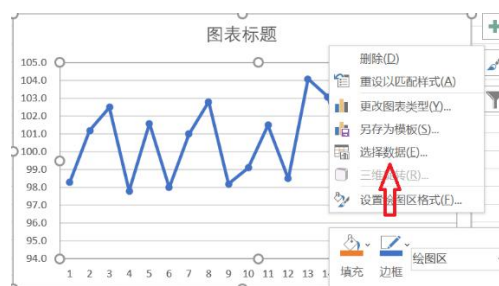
图AXX3: 图表选择

点“确定”得“图AXX4: 折线图”:



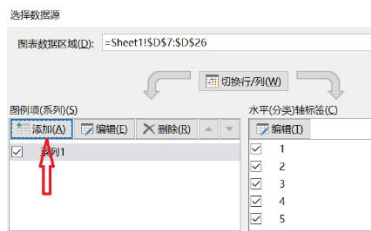
图AXX4: 折线图

将“表AXX1, 质控样数据”的平均值100.5在Excel表上复制20个, 使其数量与质控样的数量一致, 然后将鼠标放在“图AXX4: 折线图”上, 点击右键, 再点“选择数据”, 见



图AXX5: 选择数据

点“添加”, 见“图AXX6: 选择数据源”。



图AXX6：选择数据源

在“系列名称”中输入平均值，在系列值中将20个平均值在EXCEL中的区域输入，见“图AXX7：编辑数据系列”：



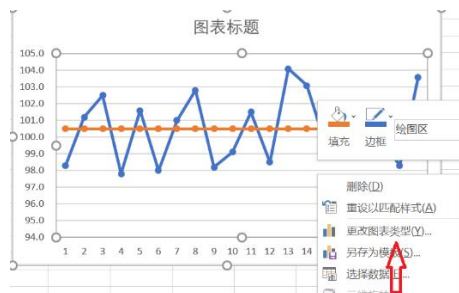
图AXX7：编辑数据系列

点“确定”，返回“选择数据源”图，再点“确定”得“图AXX8：含有平均值的折线图”：



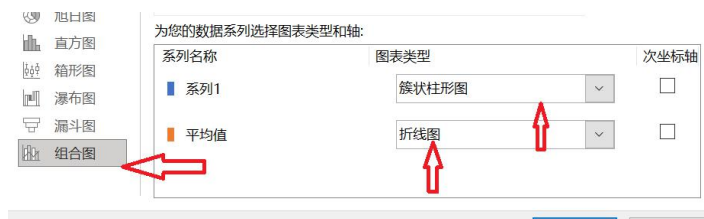
图AXX8：含有平均值的折线图

然后将鼠标放在“图AXX8：含有平均值的折线图”上，点击右键，再点“更改图表类型”，见图AXX9：修改图表类型



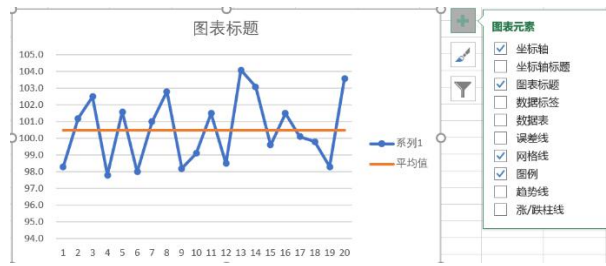
图AXX9：更改图表类型

点“组合图”，将“系列1”的图表类型改回“带数据标记的折线图”、将“平均值”改成“折线图”，见“图AXX10：组合图中更改图表类型”：



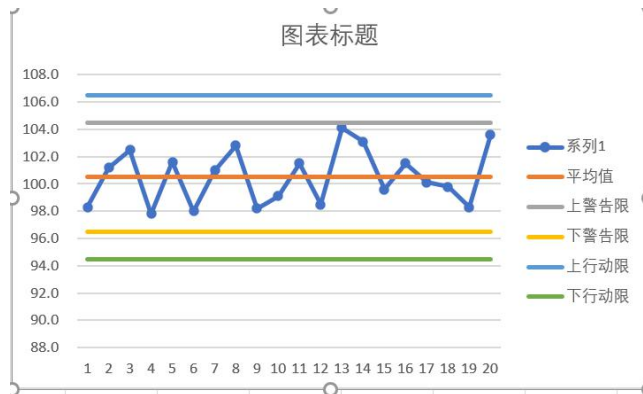
图AXX10：组合图中更改图表类型

点“确定”出现含平均值均线的控制图，点右上角“+”，在展开的“图表元素”中点“图例”，出现“图AXX11：含平均值的控制图”：



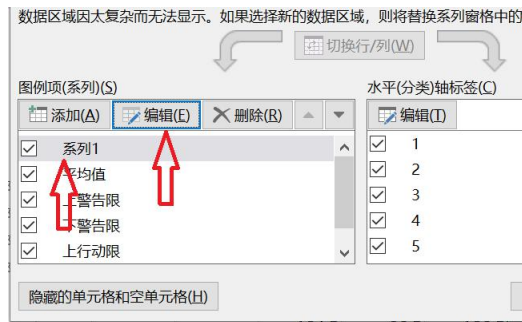
图AXX11：含平均值的控制图

将上、下警告限、上、下行动限按平均值的方式添加，得“图AXX11：含控制限的图”：



图AXX12：含控制限的图

图中“系列1”没有命名，如需要将其改成“数据”，将鼠标放在“图AXX11：含控制限的图”上，点击右键，点“选择数据”，在出来的“选择数据源”图中，选择“系列1”再选择“编辑”，见“图AXX12：编辑数据”



图AXX13：编辑数据

在出现的“编辑数据系列”图中，在“系列名称”中输入“数据”，见“图AXX14：更改系列名称”



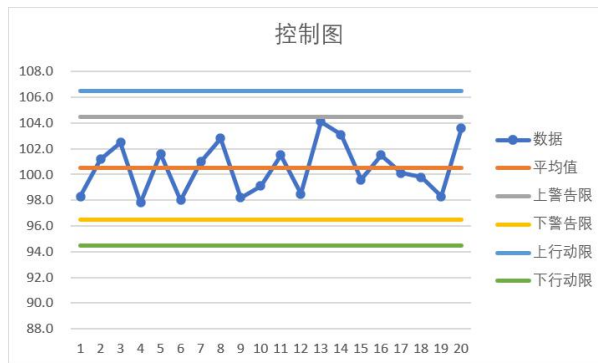
图AXX14：更改系列名称

点“确定”得到“图AXX15：更改系列名后的控制图”：



图AXX15：更改系列名后的控制图

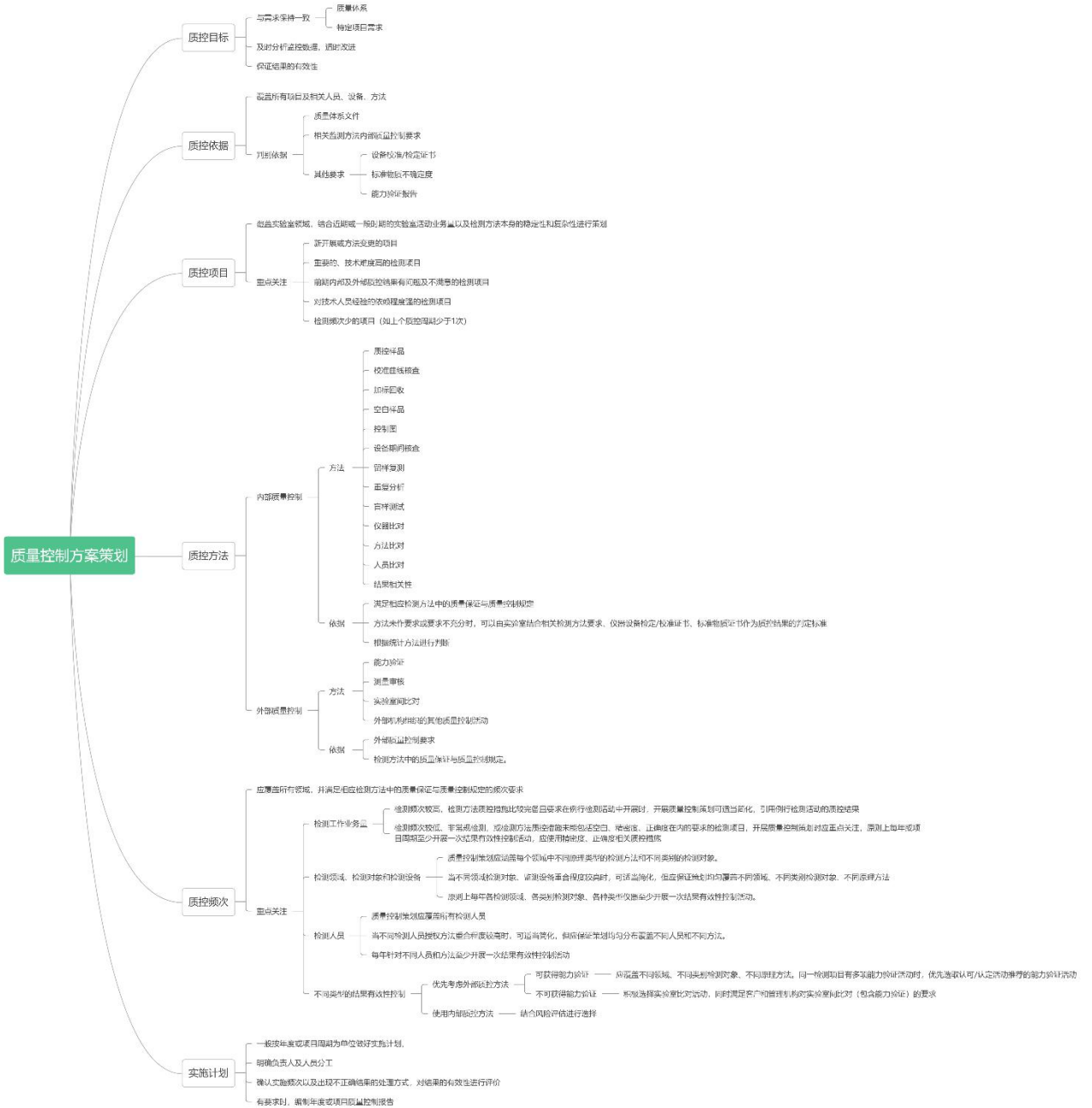
如需要将“图表标题”改为“控制图”，直接在“图AXX15：更改系列名后的控制图”上修改，最终的控制图见“图AXX16：控制图”：



图AXX16：控制图

### 附录 C (资料性)

## 结果有效性监控活动策划思维导图示例





附录 D  
(资料性)  
监控策划示意表

序号	检测对象	项目/参数		检测标准(方法)名称及编号(含年号)	检测频率	检测人员数量	主要设备		PT	新项目	常规	质控图	方法质控要求	监控频次	监控方式		判定标准		监控依据	
		序号	名称				设备	数量							检测时	例行监控	检测时	例行监控	检测时	例行监控
1	环境空气和废气	1	四到八氯代二噁英和呋喃	《环境大气中多氯代,多溴代,溴代/氯代二噁英和呋喃的检测》US EPA TO-9A(1999)	每两月1次	3	高分辨气-质联用	1	否	否	否	否	有	每月1次	内标回收率 检出限 确认 空白实验 平行实验(具备条件时)	CRM 或方法比对 或人员比对	方法	En值	方法	客户要求、行业质量规范、作业指导书
				《环境空气和废气 二噁英类的测定 同位素稀释高分辨气相色谱-高分辨质谱法》HJ77.2-2008	每两月1次	3	高分辨气-质联用	1	否	否	是	否	有							
		2	化学需氧量	水质 化学需氧量的测定 重铬酸盐法 HJ828-2017	每天	4	棕色酸式滴定管	2	是	否	是	是	有	每天	空白实验 有证标准物质 平行实验	无	方法	无	方法	无

		3	氯化物	水质 无机阴离子 (F <sup>-</sup> 、Cl <sup>-</sup> 、NO <sub>2</sub> <sup>-</sup> 、Br <sup>-</sup> 、NO <sub>3</sub> <sup>-</sup> 、PO <sub>4</sub> <sup>3-</sup> 、SO <sub>3</sub> <sup>2-</sup> 、SO <sub>4</sub> <sup>2-</sup> ) 的测定 离子色谱法 HJ84-2016	每天	4	离子色谱仪	2	是	否	是	是	有	每天	空白实验 标曲核 查 加标回 收 平行实 验	无		无	方法	无
2	铁矿石	1	水分	铁矿石 水分含量的测定 重量法 GB/T 6730.2-2018	每月4次	4	电子天平	2	否	否	是	否	无	空白实验 加标回 收 平行实 验	方法比对 /设备比 对/人员 比对/实 验室间比 对/质控 图	相对偏差 允许限	En 值/ 测 试 质 量 参 数 分 数	客户 要求 或行 业质 量规 范	客户 要求 或行 业质 量规 范	
				铁矿石 分析样中吸湿水分的测定 重量法、卡尔费休法和质量损失法 GB/T 6730.3-2017	每月4次	4	电子天平	2	否	否	是	否	无							

		2	金属铁	铁矿石 金属铁含量的测定 磺基水杨酸分光光度法 GB/T 6730.7-2016	每月4次	4	紫外可见分光光度计	1	是	否	否	是	无	空白实验 有证标准物质 平行实验	设备比对 /人员比对/实验室间比对 /质控图	重复分析 相对偏差 允许限系数 1.00	En 值/ 测试 质量 参数 分数	客户 要求 或行业 质量规 范	客户要求 或行业质 量规范
		3	硫	铁矿石 碳和硫含量的测定 高频燃烧红外吸收法 GB/T 6730.61-2005	每月4次	4	碳硫分析仪	2	是	是	否	是	无	空白实验 有证标准物质 平行实验	设备比对 /人员比对/实验室间比对 /质控图	重复分析 相对偏差 允许限系数 0.637	En 值/ 测试 质量 参数 分数	客户 要求 或行业 质量规 范	客户要求 或行业质 量规范

		4	铜	铁矿石 铜含量的测定 火焰原子吸收光谱法 GB/T 6730.36-2016	每月4次	4	原子吸收分光光度计	2	是	是	否	是	无	空白实验 有证标准物质 平行实验	设备比对 /人员比对/实验室间比对 /质控图	重复分析 相对偏差 允许限系数 1.00	En 值/ 测试 质量 参数 分数	客户要求 或行业 质量 规范	客户要求 或行业 质量 规范
--	--	---	---	--	------	---	-----------	---	---	---	---	---	---	------------------------	------------------------------	----------------------------	----------------------------------	-------------------------	-------------------------