ICS 03.120.30
A 41

中华人民共和国国家标准

GB/T 6379.3—2012/ISO 5725-3:1994

测量方法与结果的准确度(正确度与精密度) 第3部分:标准测量方法精密度的中间度量

Accuracy(trueness and precision) of measurement methods and results—Part 3:Intermediate measures of the precision of a standard measurement method

(ISO 5725-3:1994,IDT)

2012-11-05 发布　　　　2013-02-15 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会　发布

目　次

前　言

GB/T 6379《测量方法与结果的准确度(正确度与精密度)》分为以下几个部分,其预期结构及对应的国际标准为:

——第 1 部分:总则与定义(ISO 5725-1:1994,IDT)

——第 2 部分:确定标准测量方法的重复性和再现性的基本方法(ISO 5725-2:1994,IDT)

——第 3 部分:标准测量方法精密度的中间度量(ISO 5725-3:1994,IDT)

——第 4 部分:确定标准测量方法正确度的基本方法(ISO 5725-4:1994,IDT)

——第 5 部分:确定标准测量方法精密度的可替代方法 (ISO 5725-5:1998,IDT)

——第 6 部分:准确度值的实际应用(ISO 5725-6:1994,IDT)

本部分为 GB/T 6379 的第 3 部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

本部分等同采用国际标准 ISO 5725-3:1994《测量方法与结果的准确度(正确度与精密度)　第 3 部分:标准测量方法精密度的中间度量》及 ISO 于 2001-10-15 发布的对 1994 版 ISO 5725-3 的技术修改单。对 ISO 5725-3:1994 的错误作了如下的修改和更正:

——修改了 9.4 中关于根据完全套设计试验,对重复性标准差、再现性标准差及中间精密度标准差估计的不准确叙述。

——将附录 D 的表 D.5 中 $s_{I(T)}$ 在第 6 水平的值由 9.545×10^{-3} 更正为 8.020×10^{-3}。

GB/T 6379 第 1 部分至第 6 部分作为一个整体代替 GB/T 6379—1986 和 GB/T 11792—1989。标准中将原精密度概念加以扩展,增加了正确度概念,统称为准确度;除重复性条件和再现性条件外,增加了中间精密度条件。

本部分由全国统计方法应用标准化技术委员会(SAC/TC 21)提出并归口。

本部分起草单位:中国标准化研究院、中国科学院数学与系统科学研究院、深圳市华测检测有限公司、海南省产品质量监督检验所、无锡市产品质量监督检验所、广州出入境检验检疫局。

本部分主要起草人:于振凡、冯士雍、丁文兴、朱平、黄艳、陈华英、吴建国、李成明。

本部分于 2012 年首次发布。

引　言

0.1　GB/T 6379 用两个术语"正确度"与"精密度"来描述一种测量方法的准确度。正确度指大量测试结果的(算术)平均值与真值或接受参照值之间的一致程度;而精密度指测试结果之间的一致程度。

0.2　GB/T 6379.1 中对上述诸量给出了一般性考虑,在本部分中不再重复。必须强调指出,GB/T 6379.1 应与 GB/T 6379 所有其他部分(包括本部分)结合起来读,因为 GB/T 6379.1 给出了基本定义和总则。

0.3　很多不同的因素(除假定相同的样品之间的差异外)都能够引起测量方法的结果变异,这些因素包括:

a)　操作员;

b)　使用的设备;

c)　设备的校准;

d)　环境(温度、湿度、空气污染等);

e)　试剂的批;

f)　不同测量的时间间隔。

由不同操作员所做的测量和在不同设备上进行的测量通常要比在短时间内由同一个操作员使用相同的设备进行测量产生的变异大。

0.4　精密度的两个条件,即重复性条件和再现性条件是必需的,并且在许多实际情况下,对描述测量方法的变异是有用的。在重复性条件下,0.3 中所列的因素 a)～f)皆保持不变,不产生变异,而在再现性条件下,这些条件都是变化的,能引起测试结果的变异。因此,重复性条件和再现性条件是精密度条件的两个极端情形。前者描述测试结果最小变异,而后者描述测试结果最大变异。在这两种极端条件之间的中间条件也是存在的,即因素 a)～f)之中的一个或多个发生变化,它们可用于某些特定的环境。

精密度通常用标准差表示。

0.5　本部分主要讨论一种测量方法的中间精密度的度量。由于这些度量的值处于该测量方法两种极端精密度度量值(重复性标准差和再现性标准差)之间,故称中间精密度。

为了说明这些中间精密度度量的必要性,考虑当前一个与生产车间有关的实验室的运作,在这个实验室内实行三班倒的工作制,测量由不同的操作员在不同的设备上进行。因而操作员和设备是能引起测试结果变异的部分因素。在评定测量方法的精密度时必须予以考虑。

0.6　本部分定义的中间精密度度量首先应用于下面情形:为了在实验室内部对测量方法进行改进、标准化和控制,要进行中间精密度的估计;在一个特殊设计的实验室间研究中也需要进行中间精密度的估计。但是,由于 1.3 和 9.1 所述的原因,对这些度量的解释和应用应当谨慎。

0.7　最有可能影响测量方法精密度的四个因素是:

a)　时间:连续性测量的时间间隔是大还是小。

b)　校准:在连续的几组测量之间同一设备是否经过重新校准。

c)　操作员:连续的测量是否由同一个操作员完成。

d)　设备:在测量中是否使用同一设备(或同一批试剂)。

0.8　下面,先引进 M 个因素不同的中间精密度条件($M=1,2,3$ 或 4),以便考虑实验室内测量条件(时间、校准、操作员和设备)的变化。

a)　$M=1$:四个因素中只有一个不同;

b)　$M=2$:四个因素中有两个不同;

c)　$M=3$:四个因素中有三个不同;

d)　$M=4$:所有四个因素都不同。

不同的中间精密度条件产生不同的中间精密度标准差,记作 $s_{I(\cdot)}$,所对应的特定条件在圆括号里明确标出,例如,$s_{I(TO)}$ 表示不同时间、不同操作员的中间精密度标准差。

0.9　对于中间精密度条件下的测量,0.7 中所列出因素中有一个或多个不同,在重复性条件下,那些因素被假定为常量。

在重复性条件下所得测试结果的标准差,一般要小于在中间精密度条件下所得测试结果的标准差。一般情况下,在化学分析中,中间精密度条件下的标准差会是重复性条件下标准差的 2～3 倍。当然,它不应大于再现性标准差。

例如,在一个确定铜矿石中铜含量的共有 35 个实验室参与的协同试验中,发现不论使用电解比重测定法还是用 $Na_2S_2O_3$ 滴定方法,在一个因素不同(时间不同但操作员和设备相同)的中间精密度条件下,所得标准差比重复性条件下标准差大 1.5 倍。

测量方法与结果的准确度(正确度与精密度)
第3部分:标准测量方法精密度的中间度量

1 范围

1.1 本部分规定了由于实验室内观测条件(时间、校准、操作员和设备)变化而产生的四种中间精密度度量。这些中间度量可以在一个确定的实验室内试验中产生,也可以通过实验室间试验产生。此外,GB/T 6379 本部分:

a) 讨论中间精密度度量定义的含义;
b) 为在实际工作中对中间精密度度量估计的解释和应用提供指南;
c) 没有为估计中间精密度度量的误差提供任何度量;
d) 不涉及如何确定测量方法本身的正确度,但讨论了正确度与测量条件之间的关系。

1.2 本部分适用于所涉及的测量方法特指对连续量进行测量,并且每次测量只取一个值作为测量结果,尽管这个值可能是一组观测值的计算结果。

1.3 确定这些中间精密度度量的本质在于,用数量表示测量方法在规定条件下,重复测试结果的能力。

1.4 本部分所述的统计方法基于如下的前提:可以联合"相似"的测量条件中的信息,以获得对中间精密度度量更为准确的信息。只要所称的"相似"确实"相似",这个前提即是有效的。但通过实验室间研究来估计中间精密度度量时,这个前提很难得到满足。例如,为使联合不同实验室的信息有意义,需要通过控制所有参与试验的实验室的"时间"影响(效应)或"操作员"影响(效应),使它们"相似",就非常困难。因此,在使用中间精密度实验室间研究所得的结果时要加以小心。实验室内研究也依赖于上述前提,但此时由于分析者对一个因素的实际影响了解更多,也知道该如何对它进行控制,因而这个前提更易于实现。

1.5 除本部分所述的技术外,还有另外一些估计和证实一个实验室内中间精密度度量的技术,例如控制图(见 GB/T 6379.6)。本部分并未声明提供了在某一特定实验室内对中间精密度度量进行估计的唯一方法。

注:本部分涉及试验设计,例如套设计的知识。附录 B 和附录 C 中给出了相关的基础知识。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是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 3358.1—2009 统计学词汇及符号 第1部分:一般统计术语与用于概率的术语(ISO 3534-1:2006,IDT)

GB/T 6379.1—2004 测量方法与结果的准确度(正确度与精密度) 第1部分:总则与定义(ISO 5725-1:1994,IDT)

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:确定标准测量方法重复性和再现性的基本方法(ISO 5725-2:1994,IDT)

ISO 3534-1:1993 统计学 词汇和符号 第1部分:概率和一般统计术语

ISO 指南 33:1989 有证标准物料的使用

ISO 指南 35:1989 标准物料的定值 总则和统计原理

3 术语和定义

GB/T 6379.1 和 ISO 3534-1 界定的以及下列术语和定义适用于本文件。

GB/T 6379 使用的符号在附录 A 中给出。

4 一般要求

为保证测量方法的一致性,应使用标准化的测量方法。构成一个特定实验室内试验或实验室间试验一部分的所有测量都应按标准方法进行。

5 重要因素

5.1 实验室内测量条件的四个因素(时间、校准、操作者和设备)被认为是产生测量结果变异的主要原因(见表 1)。

表 1 四个重要因素及其状态

因素	实验室内的测量条件	
	状态 1(相同)	状态 2(不同)
时间	在相同时间进行的测量	在不同时间进行的测量
校准	两次测量之间不进行校准	两次测量之间进行校准
操作员	相同的操作员	不同的操作员
设备	未经重新校准的相同设备	不同的设备

5.2 "同时间测量"包括那些在尽可能短的时间内进行的测量,其目的是使试验条件(例如不能保证恒定的环境条件)的变化最小。"不同时间测量"是指那些在较长的时间间隔内进行的测量,可能由于环境条件的变化而对测量发生影响。

5.3 "校准"在此处不是指由测量方法所规定的作为获取测试结果程序中的一个组成部分的校准,而是指在一个实验室内部不同组测量之间的每隔一定时间所进行的校准过程。

5.4 对于某些操作,"操作员"事实上可能指一组操作员,每一操作员执行测量程序的某一规定部分。在此情况,"操作员"是指这一组操作员,这一组操作员中出现的任何人员或所分配任务的变更都应看作是不同的"操作员"。

5.5 "设备"事实上往往是指成套的设备。而成套设备中任何重要部件的任何变化都将被视为不同的"设备"。至于什么是重要部件,可照常识判断。温度计的变更将被视作不同的重要部件;而用一个稍微不同的容器来代替水槽将被视为无关紧要。使用不同批次的试剂应被视作重要部件变化,这将被认为是使用了不同的"设备";如果这一变化发生在某次校准之后,则被看作为一次重新校准。

5.6 在重复性条件下,所有的四个因素都处于表 1 中的状态 1。对于中间精密度条件,一个或者多个因素处于表 1 中的状态 2,称为"M 个因素不同的精密度条件",其中 M 为处于状态 2 的因素个数。在再现性条件下,测量结果由不同的实验室获得,因此不仅四个因素都处于状态 2,且由于不同实验室在实验室管理与维持、操作员的总体训练水平、测试结果的稳定性和核查等方面的不同,还会有额外的影响。

5.7 对 M 个因素不同的中间精密度条件,有必要指明哪些因素处于表 1 中的状态 2,且用相应的下标表示。例如:

——时间不同的中间精密度标准差,$S_{I(T)}$;
——校准不同的中间精密度标准差,$S_{I(C)}$;
——操作员不同的中间精密度标准差,$S_{I(O)}$;
——时间与操作员不同的中间精密度标准差,$S_{I(TO)}$;
——时间、操作员与设备不同的中间精密度标准差,$S_{I(TOE)}$;

其他情形也用类似的表示方法。

6 统计模型

6.1 基本模型

为估计测量方法的准确度(正确度和精密度),假定每个测试结果 y 是以下 3 个分量的和:

$$y=m+B+e \qquad \cdots\cdots(1)$$

其中,对给定的受试物料:

m ——总平均值(期望);
B ——重复性条件下偏倚的实验室分量;
e ——重复性条件下每次测量产生的随机误差。

以下分别讨论模型中的每一分量以及基本模型的推广。

6.2 总平均值 m

6.2.1 总平均值 m 是所有测试结果总的平均值。在一项协同研究(见 GB/T 6379.2)中获得的 m 值仅依赖于"真值"和测量方法,而不依赖于获得这些测试结果的实验室、设备、操作员和时间因素。一种特定的受试物料的总平均值称为"测试水平";例如一种化学品的不同纯度的样品或不同物料(例如不同型号的钢材)对应着不同的水平。

在许多情形,受试特性的真值 μ 的概念是适用的,例如,一种正在滴定溶液的真实浓度。水平 m 并不总是与真值 μ 相等;差值 $m-\mu$ 称为"测量方法的偏倚"。

在某些情况下,测试水平完全取决于所用的测量方法,此时一个独立的真值概念不再适用。例如,钢材的维氏(Vicker)硬度和焦炭的米库姆(Micum)转鼓指数就属于这类情况。通常用 δ 表示偏倚(真值不存在时,$\delta=0$),总平均值 m 即可表示为:

$$m=\mu+\delta \qquad \cdots\cdots(2)$$

注:对偏倚项 δ 的讨论及关于正确度试验的描述在 GB/T 6379.4 中给出。

6.2.2 在检查用相同测量方法获得的测试结果间的差异时,测量方法的偏倚不会对其产生影响,因此可以忽略,除非它依赖于测试水平。当把测试结果和一个合同中的规定值或标准值进行比较,而合同中的规定值或标准值指的是真值 μ 而不是测试水平 m 时,以及比较由不同测量方法得到的测试结果时,必须考虑测量方法的偏倚。

6.3 分量 B

6.3.1 分量 B 代表由于种种原因造成的关于 m 的实验室偏倚,它与在每一次测试中都会发生的随机误差 e 无关。在一个实验室内重复性条件下,B 被看作一个常数并被称作"偏倚的实验室分量"。

6.3.2 然而,当常规使用某种测量方法时,实验室偏倚 B 的数值中,显然包含多种效应,比如说,由操作员、所使用的设备、设备的校准,以及环境(温度、湿度、空气质量等等)的变化所产生的效应。这样,式(1)的统计模型可以改写为:

$$y=m+B_0+B_{(1)}+B_{(2)}+\cdots+e \qquad \cdots\cdots(3)$$

或

$$y = \mu + \delta + B_0 + B_{(1)} + B_{(2)} + \cdots + e \qquad \cdots\cdots(4)$$

其中 B 由 B_0，$B_{(1)}$，$B_{(2)}$ 等分量构成，由多种中间精密度因素说明。

在实际中，研究目标及测量方法的灵敏度决定了模型的复杂程度。在许多情况下，模型的简化形式就已足够。

6.4 分量 B_0，$B_{(1)}$，$B_{(2)}$ 等

6.4.1 这些分量在重复性条件下都为常量，是测试结果偏倚的一部分。在中间精密度条件下，B_0 是保持相同(表 1 中的状态 1)的诸因素的固定效应；而 $B_{(1)}$，$B_{(2)}$ 等为变化的(表 1 中的状态 2)诸因素的随机效应。它们并不增加偏倚，但增加中间精密度标准差，使之比重复性标准差要大。

6.4.2 操作员效应是由于不同操作员之间的差异产生的，包括测量方法操作中(譬如阅读标尺刻度)的个人习惯。部分此类差异应通过对测量方法的标准化，特别是提供清楚准确的技术说明予以消除。即使同一操作员得到的测试结果，偏倚的这一分量也不总是常数(例如偏倚的大小会随着操作员当日的精神或体力状况不同而变化)，而且这部分偏倚不能完全校正。应通过使用表述清楚的操作手册和培训来减少这种偏倚。在这种条件下，由操作员不同引起的效应可看作是随机的。

6.4.3 设备效应是由于不同设备之间的差异产生的，它包括由于设备安装位置的不同，特别是由于指针或记录仪等的波动产生的效应。某些这类效应能通过精确的校准来消除。因设备间系统原因产生的差异应通过校准来纠正，标准测量方法中应该包括这种校准程序，如更换一批试剂即可按上述方式处理。在执行这一程序时，需要一个接受参照值，这方面应参考 ISO 指南 33 和 ISO 指南 35。已经过标准物料校准的设备带来的剩余效应将认为是随机效应。

6.4.4 时间效应是由于不同时间的环境差异(例如室温、湿度等的变化)产生的。应通过对环境条件的标准化尽可能将这种效应降到最低。

6.4.5 一个操作员的技术或疲劳度产生的效应可以看作为操作员与时间的交互效应。一套设备的性能可能会在刚开始使用时和使用许多小时后有差异，这是设备与时间的交互效应的一个例子。当操作员人数很少而设备的数量更少时，这些因素产生的效应可看作为固定(非随机)效应。

6.4.6 GB/T 6379.2 中给出的方法的前提是假定偏倚的实验室分量是近似服从正态分布的。但实际上对大多数分布，只要是单峰的，方法都可用。B 的方差称作"实验室间方差"，表示为：

$$Var(B) = \sigma_L^2 \qquad \cdots\cdots(5)$$

然而，这一方差仍包括由于操作员、设备、时间及环境变化而产生的效应。在一个由不同操作员、在不同测量时间及不同的环境等条件下进行的精密度试验中，利用套设计，可以计算中间精密度方差。$Var(B)$ 看作由实验室、操作员、试验日期、环境等独立方差分量组成：

$$Var(B) = Var(B_0) + Var(B_{(1)}) + Var(B_{(2)}) + \cdots \qquad \cdots\cdots(6)$$

上述方差记为：

$$Var(B_0) = \sigma_{(0)}^2$$
$$Var(B_{(1)}) = \sigma_{(1)}^2$$
$$Var(B_{(2)}) = \sigma_{(2)}^2$$
$$\cdots\cdots \qquad \cdots\cdots(7)$$

在实际中，$Var(B)$ 用 S_L^2 进行估计，类似的中间精密度估计可以通过适当设计的试验得到。

6.5 误差项 e

6.5.1 误差项表示每一测试结果中都包含的随机误差。在本部分给出的方法中始终假定此误差变量近似服从正态分布。但是实际上，只要分布是单峰的即可。

6.5.2 在一个单个实验室内，误差方差称作为试验室内方差，表示为：

$$Var(e) = \sigma_W^2 \qquad \cdots\cdots(8)$$

6.5.3 由于诸如实验室间操作员技术上的差异，可以预计不同实验室的 σ_W^2 值会有所不同。但在本部分中，假定通过合理的测量方法的标准化，不同实验室的实验室内方差之间的差异很小，有理由对所有使用标准测量方法的实验室确定一个实验室内方差的公共值。这个公共值以各实验室内方差的平均值作为估计值，称作"重复性方差"，表示为：

$$\sigma_r^2 = \overline{Var(e)} \qquad \cdots\cdots(9)$$

该平均值是对离群实验室以外的所有参与准确度试验的实验室求得的。

7 测量条件的选择

7.1 在用一个测量方法进行测量时，在一个实验室内，可想象有许多测量条件，例如：

a) 重复性条件(四因素均为常量)；

b) 一个因素不同的几种中间精密度条件；

c) 两个因素不同的几种中间精密度条件；

d) 三个因素不同的几种中间精密度条件；

e) 四个因素都不同的中间精密度条件。

在测量方法标准中，尽管总应给出重复性标准差，但不必(尽管可行)给出所有可能精密度度量。在一般的商业实践中，对中间精密度测量，应说明通常会遇到的测量条件。在详细说明与之相应的特定测量条件的同时，只明确一个适当的中间精密度度量就已足够。应该仔细地说明可能改变的测量条件因素。特别对于时间不同的中间精密度，应明确连续测量间的实际平均时间间隔。

7.2 我们假定，一个标准化测量方法的偏倚已经尽可能小，且测量方法本身固有的偏倚也已通过技术手段进行过处理。因此，本部分只讨论源于测量条件的偏倚。

7.3 重复性条件中测量条件诸因素(时间、校准、操作员和设备)的任何变化(即由表 1 中的状态 1 变到状态 2)都会增大测试结果的变异。然而，多个测试结果平均值的期望的偏倚会比重复性条件下的偏倚小。因此不用单个测试结果，而用多个测试结果的平均值作为最终报告的测试结果，即能克服中间精密度条件下标准差的增大。

7.4 大多数实验室的一些实际考虑，例如所要求的最终上报结果的精密度(标准差)及进行测量的费用，将决定测量方法标准化中所考虑的测量条件改变的因素个数和因素的选择。

8 中间精密度度量的实验室内研究和分析

8.1 最简单的方法

估计一个实验室内中间精密度标准差的最简单的方法是：抽取一个样本(或对于破坏性的测试，抽取一组假定为完全相同，也即同一的样本)，对其进行次数为 n 的重复测量，在不同次测量之间因素发生改变。建议测量次数 n 至少应为 15。对实验室而言，此要求可能不能满足，且与其他方法比较，这种估计实验室内中间精密度的方法的效率也不高。然而，它的分析过程很简单，这种方法对于通过连续多日对同一样本进行连续测量来研究不同时间的中间精密度，或研究不同校准对测量的影响是有用的。

为检测数据中潜在的离群值，推荐使用 $y_k - \overline{y}$ 对测量数 k 的图，其中 y_k 是 n 个重复测试结果中第 k 个测试结果，而 $\overline{y}$ 是这 n 个重复测试结果的平均值。更正规的检测离群值的方法包括 GB/T 6379.2—2004 中 7.3.4 给出的格拉布斯(Grubbs)检验。

M 个因素不同时，中间精密度标准差的估计值由下式给出：

$$s_{I(\;)} = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(y_k - \overline{y})^2} \qquad \cdots\cdots(10)$$

表示中间精密度条件的符号应该标在下标的括号内。

8.2 可供选择的方法

8.2.1 另一种可供选择的方法要考虑 t 组测量，每组测量包括 n 个重复测试结果。例如，在一个实验室内，一组共 t 种物料，每一种经过测量后，改变中间精密度因素，对 t 种物料进行重新测量，重复这种程序直到每一种物料得到 n 个测试结果。每组中的 n 个测试结果应由同一个样本上测量得到(对破坏性测试，由同一组假定为同一的样本得到)。所测试的物料不必是同一的，唯一要求是 t 种物料都在同一测试水平的区间内，对每一种物料，只要 M 个因素不同的中间精密度标准差的一个值属于该区间就可认为此物料属于该测试水平区间。建议 $t(n-1)$ 的值至少为 15。

示例：一个操作员对 t 种物料中的每一种进行单个测量后，由第二个操作员重复这一过程，接着可能由第三个操作员进行测量，如此继续，此后就可以计算中间精密度的估计 $s_{I(O)}$。

8.2.2 为检测数据中潜在的离群值，推荐使用 $y_{jk} - \overline{y}_j$ 对物料数 j 的图，其中 y_{jk} 是对第 j 种物料的第 k 个测试结果，$\overline{y}_j$ 是第 j 种物料的 n 个测试结果的平均值。更正规的检测离群值的方法包括 GB/T 6379.2—2004 中 7.3.4 给出的格拉布斯(Grubbs)检验。这种检验既可以对每组结果单独检验，又可以把所有 tn 个结果作为整体进行检验。

M 个因素不同的中间精密度标准差的估计由下式出：

$$s_{I(\;)} = \sqrt{\frac{1}{t(n-1)}\sum_{j=1}^{t}\sum_{k=1}^{n}(y_{jk} - \overline{y}_j)^2} \qquad \cdots\cdots(11)$$

对 $n=2$(即每种物料有 2 个测试结果)，上述公式简化为：

$$s_{I(\;)} = \sqrt{\frac{1}{2t}\sum_{j=1}^{t}(y_{j1} - y_{j2})^2} \qquad \cdots\cdots(12)$$

8.3 测量条件对最终报告结果的影响

8.3.1 对时间、校准、操作员和设备诸因素的不同组合，$\overline{y}$ 的期望值各不相同，即使四因素中只有一个改变也是如此。这限制了对平均值的使用。在化学分析或物理测试中，$\overline{y}$ 被当作最终结果报告；对于商业原材料，这个最终报告的结果经常被用作对原材料的质量评估，并且在很大程度上影响产品的价格。

例：在煤的国际贸易中，交易量一次常超过 70 000 t，其中灰份的含量仅由 1 g 左右的测试量测定，若在合同中规定含灰量每相差 1%，每吨煤的价格相差 1.5 美元，那么在化学天平上灰的质量每相差 1 mg，相应的灰成分就差 0.1%，或者每吨煤的价格相差 0.15 美元。这批交易的总价格相差 10 500 (0.1×1.5×70 000)美元以上。

8.3.2 因此，最终报告的化学分析和物理测试的结果应当足够精确、高度可靠，尤其要是通用的和可再现的。对商业要求而言，一个只在一定的操作员、设备和时间条件下才能得到保证的最终报告结果是不够好的。

9 中间精密度度量的实验室间研究和分析

9.1 基本假定

由实验室间研究获得中间精密度估计依赖于如下假定：任何特定因素在所有实验室间的效应是相同的。例如，一个实验室更换操作员与另一个实验室更换操作员有相同的效应；而由时间因素引起的变化对各实验室都是相同的。如果此假定不成立，精密度中间度量的概念以及以下各节给出的中间精密度度量的估计技术就没有意义。必须密切注意离群值(不必须剔除的离群值)，因为这有助于检查是否偏离了可以将所有实验室获得信息联合的假定。一个检测潜在离群值的有效技术是把测量结果看作

多种因素水平或研究中涉及的各个实验室的函数，并以图形形式来描述。

9.2 最简单的方法

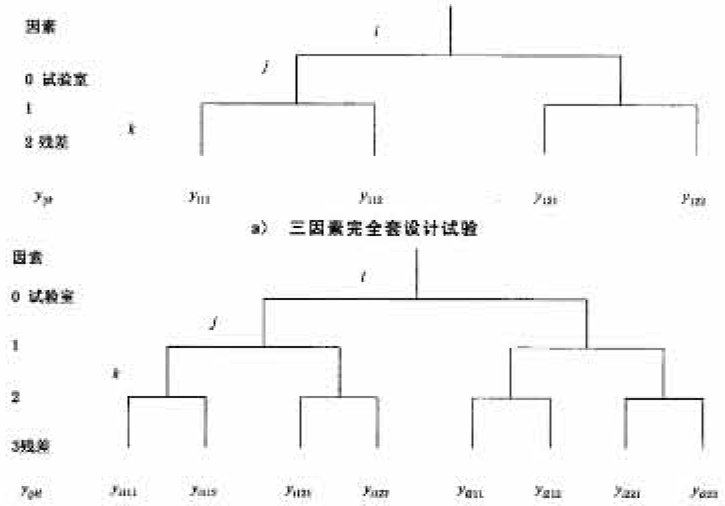
将 q 个水平的物料发送到 p 个实验室，每个实验室对 q 个水平中的每个水平进行 n 次测量，在每个水平内的 n 次测量间改变中间精密度条件。用 GB/T 6379.2 所述的同样方法进行分析，所不同的只是此时得到的是中间精密度标准差而不是重复性标准差的估计。

9.3 套设计试验

估计中间精密度的另一种方法是进行更精密复杂的试验。可能的方法有完全套设计试验和错层套设计试验(套设计定义参见 GB/T 3358.3)。使用套设计的优点是，可以通过一次实验室间试验，在同一时间，不仅获得重复性标准差和再现性标准差的估计，也同时能获得一个或者多个中间精密度标准差的估计。然而在使用套设计试验时，有些必须加以特别注意的地方，这将在 9.8 中加以解释。

9.4 完全套设计试验

图 1 是某一特定测试水平下完全套设计试验的示意图。



a) 三因素完全套设计试验

b) 四因素完全套设计试验

图 1 三因素和四因素完全套设计试验的示意图

根据几个实验室协同进行的三因素完全套设计试验，通过对 σ_r，$\sigma_{(1)}$ 及 $\sigma_{(0)}$ 的估计，可以获得重复性标准差 s_r、一个中间精密度标准差 $S_{I(1)}$ 和再现性标准差 S_R 的估计。类似地，根据四因素完全套设计试验，可以同时获得重复性标准差 S_r、两个中间精密度标准差 $S_{I(1)}$、$S_{I(2)}$ 和再现性标准差 S_R 的估计。

图 1a)中三因素完全套设计试验，数据 y 的下标 i、j 和 k，分别代表(举例说)实验室、试验日期和重复性条件下的一次重复。

图 1b)中四因素完全套设计试验，数据 y 的下标 i、j、k 和 l，分别代表(举例说)实验室、试验日期、操作员和重复性条件下的一次重复。

n 个因素完全套设计试验结果的分析是对测试的每一水平，分别采用统计中的“方差分析(ANOVA)”法进行的，附录 B 对此有详细说明。

9.5 错层套设计试验

图 2 是某一特定测试水平下错层套设计试验的示意图。

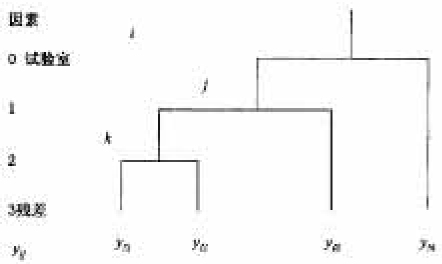


图 2 四因素错层套设计试验的示意图

三因素错层套设计试验要求每一个实验室 i 得到 3 个测试结果。测试结果 y_{i1} 和 y_{i2} 应在重复性条件下得到；y_{i3} 应在 M 个因素不同的中间精密度条件下得到($M=1,2,3$)。例如，在时间不同的中间精密度条件下的 y_{i3}，是在与得到 y_{i1} 和 y_{i2} 不同的另一日获得的。

在一个四因素错层套设计试验中，y_{i4} 应在 M 个因素不同的中间精密度条件下得到($M\geqslant 2$)。例如，在时间-操作员不同的中间精密度条件(即改变日期和操作员的条件)下得到的。

对 n 个因素错层套设计实验结果的分析是对测试的每一水平，分别采用统计中的“方差分析(ANOVA)”法进行的，附录 C 对此有详细说明。

9.6 套设计中因素的配置

套设计中因素应按如下方式安排：主要以系统效应影响的因素应放在最高层，主要以随机效应影响的因素应放在最低层，最低层的因素被看作为残差，由高到低的次序为(0，1，…)。例如，在图 1b) 及图 2 中的四因素设计试验中，因素 0 可能是实验室，因素 1 可能是操作员，因素 2 可能是测量日期，因素 3 则为重复。在完全套设计中，由于设计的对称性，因素的配置方式似乎不太重要。

9.7 套设计与 GB/T 6379.2 中给出方法的比较

GB/T 6379.2 中所给的方法，是对(测试物料的)每一测试水平分别进行分析的，实际上是两因素的完全套设计，最终获得两个标准差，即重复性标准差和再现性标准差。两个因素中，因素 0 是实验室，因素 1 是重复。如果在这一设计中增加一个因素：在一个实验室里安排两个操作员，每个操作员在重复性条件下测得两个结果，那么除了重复性标准差和再现性标准差外，还可以确定操作员不同的中间精密度标准差。如果每个实验室只用一个操作员，但是在不同的日子进行测试，就可以通过这个三因素完全套设计试验得到时间不同的中间精密度标准差。如果试验再增加一个因素，每一实验室安排两个操作员，每个操作员做两次测量并且试验在不同的工作日完全重复一次，这样安排的试验能确定重复性标准差、再现性标准差、操作员不同中间精密度标准差、时间不同中间精密度标准差以及时间-操作员不同中间精密度标准差。

9.8 完全套设计与错层套设计的比较

一个 n 因素完全套设计试验对每一实验室要求有 2^{n-1} 个测试结果，对实验室可能是一个过分的要求，这是要采用错层套设计的主要原因。尽管错层套设计的分析稍为复杂，且由于所需测试结果的数量少，标准差估计的不确定度较大，但它可以用较少的测试结果获得同样数量的标准差。

附 录 A
(规范性附录)
GB/T 6379 所用的符号与缩略语

a 关系式 $s=a+bm$ 中的截距

A 用来计算估计值的不确定度的系数

b 关系式 $s=a+bm$ 中的斜率

B 表示一个实验室测试结果与总平均值的偏差分量(偏倚的实验室分量)

B_0 表示在中间精密度条件下所有因素皆保持不变时 B 的分量

$B_{(1)}$，$B_{(2)}$，… 表示在中间精密度条件下，因素发生改变时 B 的分量

c 关系式 $\lg s=c+d\lg m$ 中的截距

C，C'，C'' 检验统计量

C_{crit}，C'_{crit}，C''_{crit} 用于统计检验的临界值

CD_P 概率 P 的临界差

CR_P 概率 P 的临界极差

d 关系式 $\lg s=c+d\lg m$ 中的斜率

e 发生在每次测试结果中随机误差分量

f 临界极差系数

$F_p(\nu_1,\nu_2)$ 自由度为 ν_1 和 ν_2 的 F 分布的 p 分位数

G 格拉布斯检验统计量

h 曼得尔实验室间一致性检验统计量

k 曼得尔实验室内一致性检验统计量

LCL 控制下限(行动限或警戒限)

m 测试特性的总平均值；水平

M 在中间精密度条件中考虑的因素数

N 交互作用数

n 一个实验室在一个水平(即一个单元中)上的测试结果数

p 参加实验室间试验的实验室数

P 概率

q 在实验室间试验中测试特性的水平数

r 重复性限

R 再现性限

RM 标准物料

s 标准差的估计值

$\hat{s}$ 标准差的预测值

T 总和

t 测试目标个数或组数

UCL 控制上限(行动限或警戒限)

W 加权回归中的权数

w 一组测试结果的极差

x 用于格拉布斯检验的数据

y 测试结果

$\bar{y}$ 测试结果的算术平均值

$\bar{\bar{y}}$ 测试结果的总平均值

α 显著性水平

β 第二类错误概率

γ 再现性标准差与重复性标准差的比值($\frac{\sigma_R}{\sigma_r}$)

Δ 实验室偏倚

$\hat{\Delta}$ Δ 的估计值

δ 测量方法偏倚

$\hat{\delta}$ δ 的估计值

λ 两个实验室偏倚或两个测量方法偏倚之间的可检出的差

μ 测试特性的真值或接受参照值

ν 自由度

ρ 方法 A 和方法 B 的重复性标准差之间的可检出的比

σ 标准差的真值

τ 表示从上次校准始由时间变化引起的测试结果变异的分量

ϕ 方法 A 和方法 B 的实验室间均方的平方根可检出的比

$\chi_p^2(\nu)$ 自由度为 ν 的 χ^2 分布的 p 分位数

用作下标的符号：

C 校准-不同

E 设备-不同

i 实验室标识

I( ) 精密度的中间度量；括号内表示中间情形类型

j 水平的标识(GB/T 6379.2)；测试或因素的标识(GB/T 6379.3)

k 实验室 i，水平为 j 的测试结果的标识

L 实验室间

m 可检出偏倚的标识

M 试样间

O 操作员-不同

r 重复性

R 再现性

T 时间-不同

W 实验室内

1,2,3,… 测试结果按获得顺序的编号

(1),(2),(3),… 测试结果按数值大小递增顺序的编号

附 录 B
(规范性附录)
完全套设计试验的方差分析

本附录中所述的方差分析必须对实验室间试验的每一测试水平分别进行。为简单起见，表明测试水平的下标没有标在测试数据上。应注意的是在本部分中，下标 j 用于表示因素 1(因素 0 代表实验室)；而在 GB/T 6379 的其他部分，j 代表测试水平。

应用 GB/T 6379.2—2004 中 7.3 描述的方法来检查数据的一致性和离群值。用本附录中所述的设计，当一个实验室的某些测试结果缺失时，对数据的准确分析将会非常复杂。如果来自某一实验室的某些测试结果被确定是岐离值或离群值，并且应在分析时予以剔除，那么，建议所有来自这一实验室的(相应水平的)数据都应在分析时予以剔除。

B.1 三因素完全套设计试验

试验中得到的数据记作 y_{ijk}，均值和极差为：

$$\bar{y}_{ij}=\frac{1}{2}(y_{ij1}+y_{ij2}) \qquad \bar{y}_i=\frac{1}{2}(\bar{y}_{i1}+\bar{y}_{i2}) \qquad \bar{\bar{y}}=\frac{1}{p}\sum_i\bar{y}_i$$

$$w_{ij(1)}=|y_{ij1}-y_{ij2}| \qquad w_{i(2)}=|\bar{y}_{i1}-\bar{y}_{i2}|$$

这里 p 为参与实验室间试验的实验室个数。

总平方和 SST 可以分解为：

$$SST=\sum_i\sum_j\sum_k(y_{ijk}-\bar{\bar{y}})^2=SS0+SS1+SSe$$

其中

$$SS0=\sum_i\sum_j\sum_k(\bar{y}_i-\bar{\bar{y}})^2=4\sum_i(\bar{y}_i-\bar{\bar{y}})^2=4\sum_i(\bar{y}_i)^2-4p(\bar{\bar{y}})^2$$

$$SS1=\sum_i\sum_j\sum_k(\bar{y}_{ij}-\bar{y}_i)^2=2\sum_i\sum_j(\bar{y}_{ij}-\bar{y}_i)^2=\sum_i w_{i(2)}^2$$

$$SSe=\sum_i\sum_j\sum_k(y_{ijk}-\bar{y}_{ij})^2=\frac{1}{2}\sum_i\sum_j w_{ij(1)}^2$$

因为平方和 SS0，SS1，SSe 的自由度分别为 $p-1$，p 和 $2p$，设计的方差分析表如表 B.1 所示。

表 B.1 三因素完全套设计试验的方差分析表

来源	平方和	自由度	均方	均方的期望
0	SS0	$p-1$	MS0= SS0/($p-1$)	$\sigma_r^2+2\sigma_{(1)}^2+4\sigma_{(0)}^2$
1	SS1	p	MS1= SS1/p	$\sigma_r^2+2\sigma_{(1)}^2$
残差	SSe	$2p$	MSe= SSe/($2p$)	σ_r^2
总和	SST	$4p-1$		

$\sigma_{(0)}^2$，$\sigma_{(1)}^2$，σ_r^2 的无偏估计值分别为 $s_{(0)}^2$，$s_{(1)}^2$ 和 s_r^2，这些估计值可由均方 MS0，MS1，MSe 按以下公式计算得到：

$$s_{(0)}^2=\frac{1}{4}(MS0-MS1)$$

$$s_{(1)}^2=\frac{1}{2}(MS1-MSe)$$

$$s_r^2= MSe$$

重复性方差、一个因素不同的中间精密度方差、再现性方差的估计值分别为：

$$s_r^2$$

$$s_{I(1)}^2=s_r^2+s_{(1)}^2$$

$$s_R^2= s_r^2+s_{(1)}^2+s_{(0)}^2$$

B.2 四因素完全套设计试验

试验中所得数据记为 y_{ijkl}，均值和极差分别为：

$$\bar{y}_{ijk}=\frac{1}{2}(y_{ijk1}+y_{ijk2}) \qquad w_{ijk(1)}=|y_{ijk1}-y_{ijk2}|$$

$$\bar{y}_{ij}=\frac{1}{2}(\bar{y}_{ij1}+\bar{y}_{ij2}) \qquad w_{ij(2)}=|\bar{y}_{ij1}-\bar{y}_{ij2}|$$

$$\bar{y}_i=\frac{1}{2}(\bar{y}_{i1}+\bar{y}_{i2}) \qquad w_{i(3)}=|\bar{y}_{i1}-\bar{y}_{i2}|$$

$$\bar{\bar{y}}=\frac{1}{p}\sum_i\bar{y}_i$$

其中 p 为参与实验室间试验的实验室个数。

总平方和 SST 可以分解为：

$$SST=\sum_i\sum_j\sum_k\sum_l(y_{ijkl}-\bar{\bar{y}})^2=SS0+SS1+SS2+SSe$$

其中

$$SS0=\sum_i\sum_j\sum_k\sum_l(\bar{y}_i-\bar{\bar{y}})^2=8\sum_i(\bar{y}_i)^2-8p(\bar{\bar{y}})^2$$

$$SS1=\sum_i\sum_j\sum_k\sum_l(\bar{y}_{ij}-\bar{y}_i)^2=4\sum_i\sum_j(\bar{y}_{ij}-\bar{y}_i)^2=2\sum_i w_{i(3)}^2$$

$$SS2=\sum_i\sum_j\sum_k\sum_l(\bar{y}_{ijk}-\bar{y}_{ij})^2=2\sum_i\sum_j\sum_k(\bar{y}_{ijk}-\bar{y}_{ij})^2=\sum_i\sum_j w_{ij(2)}^2$$

$$SSe=\sum_i\sum_j\sum_k\sum_l(y_{ijkl}-\bar{y}_{ijk})^2=\frac{1}{2}\sum_i\sum_j\sum_k w_{ijk(1)}^2$$

因为平方和 SS0，SS1，SS2，SSe 的自由度分别为 $p-1$，p，$2p$ 和 $4p$。设计的方差分析表如表 B.2 所示：

表 B.2 四因素完全套设计试验的方差分析表

来源	平方和	自由度	均方	均方的期望
0	SS0	$p-1$	MS0= SS0/($p-1$)	$\sigma_r^2+2\sigma_{(2)}^2+4\sigma_{(1)}^2+8\sigma_{(0)}^2$，
1	SS1	p	MS1= SS1/p	$\sigma_r^2+2\sigma_{(2)}^2+4\sigma_{(1)}^2$
2	SS2	$2p$	MS2= SS2/($2p$)	$\sigma_r^2+2\sigma_{(2)}^2$
残差	SSe	$4p$	MSe= SSe/($4p$)	σ_r^2
总和	SST	$8p-1$		

$\sigma_{(0)}^2$，$\sigma_{(1)}^2$，$\sigma_{(2)}^2$，σ_r^2 的无偏估计值分别为 $s_{(0)}^2$，$s_{(1)}^2$，$s_{(2)}^2$ 和 s_r^2，这些估计值可以由均方 MS0，MS1，MS2 和 MSe 按以下公式计算得到：

$$s_{(0)}^2=\frac{1}{8}(MS0-MS1)$$

$$s_{(1)}^2=\frac{1}{4}(\mathrm{MS1}-\mathrm{MS2})$$

$$s_{(2)}^2=\frac{1}{2}(\mathrm{MS2}-\mathrm{MSe})$$

$$s_r^2=\mathrm{MSe}$$

重复性方差、一个因素不同的中间精密度方差、两个因素不同的中间精密度方差及再现性方差的估计值分别为：

$$s_r^2$$

$$s_{\mathrm{I}(1)}^2=s_r^2+s_{(2)}^2$$

$$s_{\mathrm{I}(2)}^2=s_r^2+s_{(2)}^2+s_{(1)}^2$$

$$s_R^2=s_r^2+s_{(2)}^2+s_{(1)}^2+s_{(0)}^2$$

附 录 C
（规范性附录）
错层套设计试验的方差分析

本附录中所述的方差分析必须是对实验室间试验的每一测试水平分别进行的。为简单起见，表明测试水平的下标没有标在测试数据上。应注意的是在本部分，下标 j 用于表示一个实验室内的重复；而在 GB/T 6379 其他部分，j 代表测试水平。

应用 GB/T 6379.2—2004 中 7.3 描述的方法来检查数据的一致性和离群值。用本附录中所述的设计，当一个实验室的某些测试结果缺失时，对数据的准确分析将会非常复杂。如果来自某一实验室的某些测试结果被确定是歧离值或离群值，并且应在分析时予以剔除，那么，建议所有来自这一实验室的（相应水平的）数据都应在分析时予以剔除。

C.1 三因素错层套设计试验

试验中在实验室 i 得到的数据记作 $y_{ij}(j=1,2,3)$，均值和极差分别为：

$$\bar{y}_{i(1)}=\frac{1}{2}(y_{i1}+y_{i2}) \qquad w_{i(1)}=|y_{i1}-y_{i2}|$$

$$\bar{y}_{i(2)}=\frac{1}{3}(y_{i1}+y_{i2}+y_{i3}) \qquad w_{i(2)}=|\bar{y}_{i(1)}-y_{i3}|$$

$$\bar{\bar{y}}=\frac{1}{p}\sum_i \bar{y}_{i(2)}$$

其中 p 为参与实验室间试验的实验室个数。

总平方和 SST 可以分解为：

$$\mathrm{SST}=\sum_i\sum_j(y_{ij}-\bar{\bar{y}})^2=\mathrm{SS0}+\mathrm{SS1}+\mathrm{SSe}$$

其中

$$\mathrm{SS0}=3\sum_i(\bar{y}_{i(2)})^2-3p(\bar{\bar{y}})^2$$

$$\mathrm{SS1}=\frac{2}{3}\sum_i w_{i(2)}^2$$

$$\mathrm{SSe}=\frac{1}{2}\sum_i w_{i(1)}^2$$

因为平方和 SS0，SS1，SSe 的自由度分别为 $p-1$，p 和 p，设计的方差分析表如表 C.1 所示。

表 C.1 三因素错层套设计试验的方差分析表

来源	平方和	自由度	均方	均方的期望
0	SS0	$p-1$	SS0/$(p-1)$	$\sigma_r^2+\frac{5}{3}\sigma_{(1)}^2+3\sigma_{(0)}^2$，
1	SS1	p	SS1/p	$\sigma_r^2+\frac{4}{3}\sigma_{(1)}^2$
残差	SSe	p	SSe/p	σ_r^2
总和	SST	$3p-1$		

$\sigma_{(0)}^2$，$\sigma_{(1)}^2$，σ_r^2 的无偏估计值分别为 $s_{(0)}^2$，$s_{(1)}^2$ 和 s_r^2，这些估计值可由均方 MS0，MS1 和 MSe 按以下公式计算得到：

$$s_{(0)}^2=\frac{1}{3}\mathrm{MS0}-\frac{5}{12}\mathrm{MS1}+\frac{1}{12}\mathrm{MSe}$$

$$s_{(1)}^2=\frac{3}{4}\mathrm{MS1}-\frac{3}{4}\mathrm{MSe}$$

$$s_r^2=\mathrm{MSe}$$

重复性方差、一个因素不同的中间精密度方差及再现性方差的估计值分别为：

$$s_r^2$$

$$s_{\mathrm{I}(1)}^2=s_r^2+s_{(1)}^2$$

$$s_R^2=s_r^2+s_{(1)}^2+s_{(0)}^2$$

C.2 四因素错层套设计试验

试验中在实验室 i 得到的数据记作 $y_{ij}(j=1,2,3,4)$，均值和极差分别为：

$$\bar{y}_{i(1)}=\frac{1}{2}(y_{i1}+y_{i2}) \qquad w_{i(1)}=|y_{i1}-y_{i2}|$$

$$\bar{y}_{i(2)}=\frac{1}{3}(y_{i1}+y_{i2}+y_{i3}) \qquad w_{i(2)}=|\bar{y}_{i(1)}-y_{i3}|$$

$$\bar{y}_{i(3)}=\frac{1}{4}(y_{i1}+y_{i2}+y_{i3}+y_{i4}) \qquad w_{i(3)}=|\bar{y}_{i(2)}-y_{i4}|$$

$$\bar{\bar{y}}=\frac{1}{p}\sum_i \bar{y}_{i(3)}$$

其中 p 为参与实验室间试验的实验室个数。设计的方差分析表如表 C.2 所示。

表 C.2 四因素错层套设计试验的方差分析表

来源	平方和	自由度	均方	均方的期望
0	$4\sum_i(\bar{y}_{i(3)})^2-4p(\bar{\bar{y}})^2$	$p-1$	SS0/$(p-1)$	$\sigma_r^2+\frac{3}{2}\sigma_{(2)}^2+\frac{5}{2}\sigma_{(1)}^2+4\sigma_{(0)}^2$
1	$\frac{3}{4}\sum_i w_{i(3)}^2$	p	SS1/p	$\sigma_r^2+\frac{7}{6}\sigma_{(2)}^2+\frac{3}{2}\sigma_{(1)}^2$
2	$\frac{2}{3}\sum_i w_{i(2)}^2$	p	SS2/p	$\sigma_r^2+\frac{4}{3}\sigma_{(2)}^2$
残差	$\frac{1}{2}\sum_i w_{i(1)}^2$	p	SSe/p	σ_r^2
总和	$\sum_i\sum_j(y_{ij}-\bar{\bar{y}})^2$	$4p-1$		

C.3 五因素错层套设计试验

试验中在实验室 i 得到的数据记作 $y_{ij}(j=1,2,3,4,5)$，均值和极差分别为：

$\bar{y}_{i(1)}=\frac{1}{2}(y_{i1}+y_{i2})$ $\qquad w_{i(1)}=|y_{i1}-y_{i2}|$

$\bar{y}_{i(2)}=\frac{1}{3}(y_{i1}+y_{i2}+y_{i3})$ $\qquad w_{i(2)}=|\bar{y}_{i(1)}-y_{i3}|$

$\bar{y}_{i(3)}=\frac{1}{4}(y_{i1}+y_{i2}+y_{i3}+y_{i4})$ $\qquad w_{i(3)}=|\bar{y}_{i(2)}-y_{i4}|$

$\bar{y}_{i(4)}=\frac{1}{5}(y_{i1}+y_{i2}+y_{i3}+y_{i4}+y_{i5})$ $\qquad w_{i(4)}=|\bar{y}_{i(3)}-y_{i5}|$

$\bar{\bar{y}}=\frac{1}{p}\sum_i \bar{y}_{i(4)}$

其中 p 为参与实验室间协同试验的实验室个数。设计的方差分析表如表 C.3 所示。

表 C.3 五因素错层套设计试验的方差分析表

来源	平方和	自由度	均方	均方的期望
0	$5\sum_i(\bar{y}_{i(4)})^2-5p(\bar{\bar{y}})^2$	$p-1$	SS0/($p-1$)	$\sigma_r^2+\frac{7}{5}\sigma_{(3)}^2+\frac{11}{5}\sigma_{(2)}^2+\frac{17}{5}\sigma_{(1)}^2+5\sigma_{(0)}^2$
1	$\frac{4}{5}\sum_i w_{i(4)}^2$	p	SS1/p	$\sigma_r^2+\frac{11}{10}\sigma_{(3)}^2+\frac{13}{10}\sigma_{(2)}^2+\frac{8}{5}\sigma_{(1)}^2$
2	$\frac{3}{4}\sum_i w_{i(3)}^2$	p	SS2/p	$\sigma_r^2+\frac{7}{6}\sigma_{(3)}^2+\frac{3}{2}\sigma_{(2)}^2$
3	$\frac{2}{3}\sum_i w_{i(2)}^2$	p	SS3/p	$\sigma_r^2+\frac{4}{3}\sigma_{(3)}^2$
残差	$\frac{1}{2}\sum_i w_{i(1)}^2$	p	SSe/p	σ_r^2
总和	$\sum_i\sum_j(y_{ij}-\bar{\bar{y}})^2$	$5p-1$		

C.4 六因素错层套设计试验

试验中在实验室 i 得到的数据记作 y_{ij}($j=1,2,3,4,5,6$),均值和极差分别为:

$\bar{y}_{i(1)}=\frac{1}{2}(y_{i1}+y_{i2})$ $\qquad w_{i(1)}=|y_{i1}-y_{i2}|$

$\bar{y}_{i(2)}=\frac{1}{3}(y_{i1}+y_{i2}+y_{i3})$ $\qquad w_{i(2)}=|\bar{y}_{i(1)}-y_{i3}|$

$\bar{y}_{i(3)}=\frac{1}{4}(y_{i1}+y_{i2}+y_{i3}+y_{i4})$ $\qquad w_{i(3)}=|\bar{y}_{i(2)}-y_{i4}|$

$\bar{y}_{i(4)}=\frac{1}{5}(y_{i1}+y_{i2}+y_{i3}+y_{i4}+y_{i5})$ $\qquad w_{i(4)}=|\bar{y}_{i(3)}-y_{i5}|$

$\bar{y}_{i(5)}=\frac{1}{6}(y_{i1}+y_{i2}+y_{i3}+y_{i4}+y_{i5}+y_{i6})$ $\qquad w_{i(5)}=|\bar{y}_{i(4)}-y_{i6}|$

$\bar{\bar{y}}=\frac{1}{p}\sum_i \bar{y}_{i(5)}$

其中 p 为参与实验室间协同试验的实验室个数。设计的方差分析表如表 C.4 所示。

表 C.4 六因素错层套设计试验的方差分析表

来源	平方和	自由度	均方	均方的期望
0	$6\sum_i(\bar{y}_{i(4)})^2-6p(\bar{\bar{y}})^2$	$p-1$	SS0/($p-1$)	$\sigma_r^2+\frac{4}{3}\sigma_{(4)}^2+2\sigma_{(3)}^2+3\sigma_{(2)}^2+\frac{13}{3}\sigma_{(1)}^2+6\sigma_{(0)}^2$
1	$\frac{5}{6}\sum_i w_{i(5)}^2$	p	SS1/p	$\sigma_r^2+\frac{16}{15}\sigma_{(4)}^2+\frac{6}{5}\sigma_{(3)}^2+\frac{7}{5}\sigma_{(2)}^2+\frac{5}{3}\sigma_{(1)}^2$
2	$\frac{4}{5}\sum_i w_{i(4)}^2$	p	SS2/p	$\sigma_r^2+\frac{11}{10}\sigma_{(4)}^2+\frac{13}{10}\sigma_{(3)}^2+\frac{8}{5}\sigma_{(2)}^2$
3	$\frac{3}{4}\sum_i w_{i(3)}^2$	p	SS3/p	$\sigma_r^2+\frac{7}{6}\sigma_{(4)}^2+\frac{3}{2}\sigma_{(3)}^2$
4	$\frac{2}{3}\sum_i w_{i(2)}^2$	p	SS4/p	$\sigma_r^2+\frac{4}{3}\sigma_{(4)}^2$
残差	$\frac{1}{2}\sum_i w_{i(1)}^2$	p	SSe/p	σ_r^2
总和	$\sum_i\sum_j(y_{ij}-\bar{\bar{y}})^2$	$5p-1$		

附 录 D
(资料性附录)
中间精密度试验统计分析实例

D.1 例 1:在一个确定的实验室内、某一特定测试水平下得到"时间-操作员"不同的中间精密度标准差 $s_{I(TO)}$

D.1.1 背景

a) 测量方法:用真空发射光谱测定法测定钢铁中碳成分的含量,测试结果用质量百分比表示。

b) 资料来源:某钢铁厂 1984 年 11 月的常规报告。

c) 试验设计:从待测物料中随机选取的一个样本,在一个确定的实验室内,由两个分析员在前后连续的两天每天由一人对样本进行测试,按此程序,在一个月内得到 29 对这样的数据(见表 D.1)。

表 D.1 原始数据:碳含量 %(m/m)

样本号 j	第一天 y_{j1}	第二天 y_{j2}	极差 w_j	样本号 j	第一天 y_{j1}	第二天 y_{j2}	极差 w_j
1	0.130	0.127	0.003	16	0.149	0.144	0.005
2	0.140	0.132	0.008	17	0.044	0.044	0.000
3	0.078	0.080	0.002	18	0.127	0.122	0.005
4	0.110	0.113	0.003	19	0.050	0.048	0.002
5	0.126	0.128	0.002	20	0.042	0.146	0.104
6	0.036	0.032	0.004	21	0.150	0.145	0.005
7	0.050	0.047	0.003	22	0.135	0.133	0.002
8	0.143	0.140	0.003	23	0.044	0.045	0.001
9	0.091	0.089	0.002	24	0.100	0.161	0.061
10	0.040	0.030	0.010	25	0.132	0.131	0.001
11	0.110	0.113	0.003	26	0.047	0.045	0.002
12	0.142	0.145	0.003	27	0.168	0.165	0.003
13	0.143	0.150	0.007	28	0.092	0.088	0.004
14	0.169	0.165	0.004	29	0.041	0.043	0.002
15	0.169	0.173	0.004				

D.1.2 分析

数据 y_{j1},y_{j2} 和 $w_j=|y_{j1}-y_{j2}|$ 如表 D.1 所示。用 8.2 给出的方法进行分析。

图 D.1 为数据散点图[每天测试结果对测试结果均值的偏离($y_{jk}-\bar{y}_j$)对应样本号 j]。从这个散点图或利应用科克伦检验都可以查出排序在 20 和 24 位的样本是离群值。这两个样本两天的测试结果间存在着很大差异,主要原因可能是记录数据时的误差。计算时间不同、操作员不同的中间精密度标准差 $s_{I(TO)}$ 时,要剔除这两个样本的测试值。按公式(12)计算为:

$$s_{I(TO)}=\sqrt{\frac{1}{2\times 27}\sum_{j=1}^{27}w_j^2}=2.87\times 10^{-3}$$

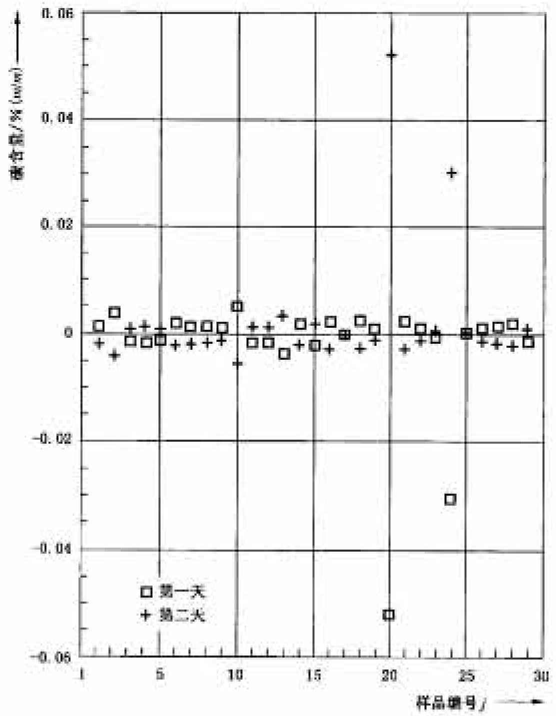


图 D.1 钢中的碳含量—每天测试结果对两天测试结果均值的偏离对应样本号

D.2 例 2 用实验室间试验得到时间不同的中间精密度标准差

D.2.1 背景

a) 测量方法：用试验说明中介绍的原子吸收光谱测定法测定钢中钒含量，测试结果用质量百分数表示。

b) 资料来源：ISO/TC 17，钢/SC 1，化学成分的测定方法和 1985 年 5 月进行的试验。

c) 试验设计：20 个实验室参与的三因素错层套设计试验中，每个实验室第一天报告两个在重复性条件下得到的测试结果，第二天在该试验所包含的 6 个水平中的每一水平下再得到并报告一个测试结果。每一实验室内的所有测量都由同一操作员、使用同一测量设备完成。

D.2.2 分析

所有 6 个水平的数据都已在表 D.2 中给出，但仅给出了其中水平 1 的方差分析。

图 D.2 给出了数据散点图(第一天和第二天的测试结果对实验室编号 i)。本图中可以看出编号 20 的实验室是离群实验室。该实验室第二天的测试结果与第一天测试结果的均值之间有一个较大的差异，这个差异与其他实验室的测试结果差异相比较是非常大的。在计算精密度度量时，剔除该实验室的测试结果。

根据附录 C 的 C.1，计算 $w_{i(1)}$，$w_{i(2)}$ 和 $\bar{y}_{i(2)}$，并在表 D.3 中给出计算结果。

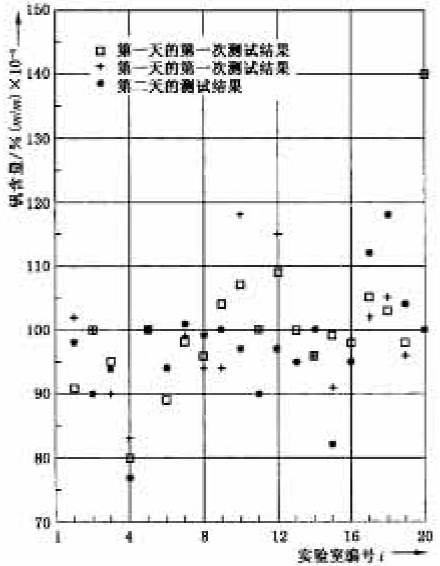


图 D.2 钢中的碳含量—水平 1 时第一天和第二天的测试结果对应实验室编号

表 D.2 钒含量的原始数据

%(m/m)

实验室号 i	水平 1(0.01%)			水平 2(0.04%)			水平 3(0.1%)		
	第一天		第二天	第一天		第二天	第一天		第二天
	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}
1	0.009 1	0.010 2	0.009 8	0.038 2	0.038 8	0.038 5	0.101	0.103	0.102
2	0.010 0	0.010 0	0.009 0	0.041 0	0.041 0	0.039 0	0.111	0.111	0.108
3	0.009 5	0.009 0	0.009 4	0.039 0	0.038 0	0.037 0	0.108	0.110	0.107
4	0.008 0	0.008 3	0.007 7	0.037 4	0.036 1	0.038 2	0.109	0.106	0.104
5	0.010 0	0.010 0	0.010 0	0.035 0	0.037 0	0.037 0	0.103	0.103	0.110
6	0.008 9	0.009 4	0.009 4	0.036 8	0.036 8	0.037 7	0.106	0.106	0.108
7	0.009 8	0.009 9	0.010 1	0.037 6	0.038 0	0.038 4	0.107	0.105	0.108
8	0.009 6	0.009 4	0.009 9	0.037 9	0.036 6	0.037 9	0.108	0.107	0.108
9	0.010 4	0.009 4	0.010 0	0.036 5	0.037 0	0.036 7	0.104	0.106	0.105
10	0.010 7	0.011 8	0.009 7	0.037 0	0.037 5	0.038 0	0.105	0.110	0.105
11	0.010 0	0.010 0	0.009 0	0.038 0	0.038 0	0.037 5	0.102	0.102	0.102
12	0.010 9	0.011 5	0.009 7	0.039 0	0.039 0	0.039 0	0.101	0.108	0.105
13	0.010 0	0.009 5	0.009 5	0.037 5	0.037 5	0.037 5	0.103	0.104	0.108
14	0.009 6	0.009 6	0.010 0	0.037 4	0.037 4	0.038 9	0.104	0.106	0.110
15	0.009 9	0.009 1	0.008 2	0.038 1	0.037 5	0.039 2	0.109	0.106	0.107
16	0.009 8	0.010 0	0.009 5	0.037 3	0.037 7	0.039 7	0.105	0.105	0.104
17	0.010 5	0.010 2	0.011 2	0.038 9	0.038 2	0.037 3	0.107	0.108	0.104
18	0.010 3	0.010 5	0.011 8	0.038 2	0.038 0	0.037 4	0.103	0.104	0.103
19	0.009 8	0.009 6	0.010 4	0.038 3	0.037 5	0.036 6	0.110	0.109	0.104
20	0.014 0	0.014 0	0.010 0	0.037 0	0.040 8	0.036 9	0.104	0.106	0.107

表 D.2（续）

%(m/m)

实验室号 i	水平 4(0.2%)			水平 5(0.5%)			水平 6(0.75%)		
	第一天		第二天	第一天		第二天	第一天		第二天
	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}
1	0.214	0.211	0.210	0.514	0.510	0.513	0.755	0.753	0.751
2	0.220	0.220	0.215	0.520	0.540	0.540	0.800	0.755	0.750
3	0.213	0.215	0.215	0.500	0.514	0.504	0.738	0.730	0.724
4	0.214	0.222	0.201	0.519	0.518	0.518	0.744	0.742	0.732
5	0.210	0.210	0.205	0.495	0.500	0.512	0.743	0.753	0.750
6	0.232	0.240	0.221	0.526	0.532	0.513	0.733	0.740	0.746
7	0.215	0.215	0.216	0.521	0.519	0.526	0.754	0.756	0.756
8	0.193	0.195	0.210	0.507	0.493	0.511	0.732	0.729	0.732
9	0.211	0.205	0.213	0.509	0.515	0.515	0.734	0.738	0.747
10	0.210	0.220	0.225	0.520	0.520	0.525	0.760	0.760	0.765
11	0.213	0.211	0.214	0.513	0.516	0.514	0.746	0.748	0.746
12	0.208	0.215	0.210	0.509	0.528	0.510	0.758	0.748	0.750
13	0.212	0.222	0.215	0.510	0.520	0.505	0.735	0.755	0.750
14	0.218	0.218	0.212	0.520	0.528	0.522	0.740	0.735	0.742
15	0.214	0.210	0.211	0.510	0.510	0.515	0.749	0.729	0.744
16	0.215	0.212	0.218	0.519	0.517	0.531	0.754	0.751	0.759
17	0.214	0.210	0.209	0.517	0.515	0.514	0.735	0.728	0.741
18	0.224	0.218	0.217	0.515	0.514	0.517	0.788	0.798	0.787
19	0.217	0.215	0.215	0.530	0.525	0.520	0.755	0.745	0.740
20	0.214	0.214	0.203	0.518	0.518	0.481	0.730	0.737	0.658

表 D.3 $w_{i(1)}$，$w_{i(2)}$ 和 $\bar{y}_{i(2)}$ 的值

实验室号 i	$w_{i(1)}$	$w_{i(2)}$	$\bar{y}_{i(2)}$
1	0.001 1	0.000 15	0.009 700
2	0.000 0	0.001 00	0.009 667
3	0.000 5	0.000 15	0.009 300
4	0.000 3	0.000 45	0.008 000
5	0.000 0	0.000 00	0.010 000
6	0.000 5	0.000 25	0.009 233
7	0.000 1	0.000 25	0.009 933
8	0.000 2	0.000 40	0.009 633
9	0.001 0	0.000 10	0.009 933
10	0.001 1	0.001 55	0.010 733
11	0.000 0	0.001 00	0.009 667
12	0.000 6	0.001 50	0.010 700
13	0.000 5	0.000 25	0.009 667
14	0.000 0	0.000 40	0.009 733
15	0.000 8	0.001 30	0.009 067
16	0.000 2	0.000 40	0.009 767
17	0.000 3	0.000 85	0.010 633
18	0.000 2	0.001 40	0.010 867
19	0.000 2	0.000 70	0.009 933

$w_{i(1)}$，$w_{i(2)}$ 和 $\bar{y}_{i(2)}$ 的平方和以及平均值 $\bar{\bar{y}}$ 的计算结果为：

$$\sum_i w_{i(1)}^2 = 5.52 \times 10^{-6}$$

$$\sum_i w_{i(2)}^2 = 12.44 \times 10^{-6}$$

$$\sum_i (\bar{y}_{i(2)})^2 = 1\,832.16 \times 10^{-6}$$

$$\bar{\bar{y}} = \frac{1}{19} \sum_i \bar{y}_{i(2)} = 0.009\,798\,25$$

利用上述结果可求得平方和 SS0，SS1 和 SSe，表 D.4 为相应的方差分析表。

分别得到不同实验室间方差的无偏估计 $s_{(0)}^2$、同一实验室、不同日期的方差的无偏估计 $s_{(1)}^2$ 以及重复性方差的无偏估计 s_r^2 为：

$$s_{(0)}^2 = 0.278 \times 10^{-6}$$

$$s_{(1)}^2 = 0.218 \times 10^{-6}$$

$$s_r^2 = 0.145 \times 10^{-6}$$

再现性标准差 s_R，时间不同的中间精密度标准差 $s_{I(T)}$ 以及重复性标准差 s_r 分别为：

$$s_R = \sqrt{s_r^2 + s_{(1)}^2 + s_{(0)}^2} = 0.801 \times 10^{-3}$$

$$s_{I(T)} = \sqrt{s_r^2 + s_{(1)}^2} = 0.603 \times 10^{-3}$$

$$s_r = \sqrt{s_r^2} = 0.381 \times 10^{-3}$$

表 D.5 为 6 个测试水平下钢中钒含量标准差的估计值的计算结果；图 D.3 为计算结果的图示。

表 D.4 钢中的钒含量的方差分析表

数据来源	平方和	自由度	均方	均方的期望
0（实验室）	24.16×10^{-6}	18	1.342×10^{-6}	$\sigma_r^2 + \frac{5}{3}\sigma_{(1)}^2 + 3\sigma_{(0)}^2$，
1（天）	8.29×10^{-6}	19	0.436×10^{-6}	$\sigma_r^2 + \frac{4}{3}\sigma_{(1)}^2$
残差	2.76×10^{-6}	19	0.145×10^{-6}	σ_r^2
总和	35.21×10^{-6}	56		

表 D.5 6 个测试水平下钢中钒含量标准差估计 s_r、$s_{I(T)}$ 和 s_R 的值

水平	离群实验室号	算术平均值/%	s_r/%	$s_{I(T)}$/%	s_R/%
1	20	0.009 8	0.381×10^{-3}	0.603×10^{-3}	0.801×10^{-3}
2	2	0.037 8	0.820×10^{-3}	0.902×10^{-3}	0.954×10^{-3}
3	—	0.105 9	1.739×10^{-3}	2.305×10^{-3}	2.650×10^{-3}
4	6 和 8	0.213 8	3.524×10^{-3}	4.710×10^{-3}	4.825×10^{-3}
5	20	0.516 4	6.237×10^{-3}	6.436×10^{-3}	9.412×10^{-3}
6	20	0.748 4	9.545×10^{-3}	8.020×10^{-3} 1)	15.962×10^{-3}

1) 在 ISO 5725-3:1994 中，表 D.5 第 6 水平的 $s_{I(T)}$ 值误为 9.545×10^{-3}，现更正为 8.020×10^{-3}（%）。

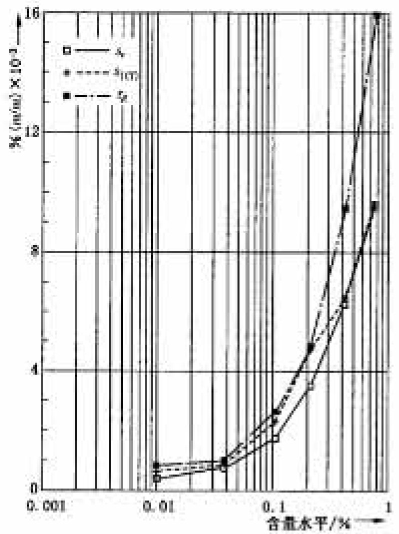


图 D.3 钢中的钒含量——重复性标准差 s_r、时间不同的中间精密度标准差 $s_{I(T)}$ 和再现性标准差 s_R 与钒成分含量水平的关系

参 考 文 献

[1] GB/T 3358.2—2009 统计学词汇及符号 第 2 部分 应用统计.

[2] GB/T 3358.3—2009 统计学词汇及符号 第 3 部分 实验设计.

[3] ISO 5725-4:1994 测量方法与结果的准确度(正确度与精密度) 第 4 部分：确定标准测量方法正确度的基本方法 Accuracy (trueness and precision) of measurement methods and results—Part 4: Basic methods for the determination of the trueness of a standard measurement method.

[4] ISO 5725-5:1998 测量方法与结果的准确度(正确度与精密度) 第 5 部分：确定标准测量方法正确度的可替代方法 Accuracy (trueness and precision) of measurement methods and results—Part 5: Alternative methods for the determination of the precision of a standard measurement method.

[5] ISO 5725-6:1994 测量方法与结果的准确度(正确度与精密度) 第 6 部分：准确度值的实际应用 Accuracy (trueness and precision) of measurement methods results—Part 6: Use in practice of accuracy values.

[6] WINER, B. J. Statistical principles in experimental design, McGraw-Hill, 1962.

[7] SNEDECOR, G. W. and COCHRAN, W. G. Statistical methods, Iowa University press. 1967.

GB/T 6379.3-2012

*

书号：155066 · 1-46246

定价： 30.00 元