ICS 03.120.30
A 41

中华人民共和国国家标准

GB/T 6379.5—2006/ISO 5725-5:1998
部分代替 GB/T 6379—1986
GB/T 11792—1989

测量方法与结果的准确度(正确度与精密度) 第5部分:确定标准测量方法精密度的可替代方法

Accuracy (trueness and precision) of measurement methods and results—Part 5: alternative methods for the determination of the precision of a standard measurement method

(ISO 5725-5:1998,IDT)

2006-11-13 发布　　2007-04-01 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 …… Ⅲ
引言 …… Ⅳ
1　范围 …… 1
2　规范性引用文件 …… 1
3　定义 …… 1
4　分割水平设计 …… 1
4.1　分割水平设计的应用 …… 1
4.2　分割水平设计安排 …… 2
4.3　分割水平试验的组织 …… 2
4.4　统计模型 …… 3
4.5　分割水平试验数据的统计分析 …… 4
4.6　对数据一致性与离群值的检查 …… 5
4.7　报告分割水平试验的结果 …… 5
4.8　例1:分割水平试验——蛋白质的测定 …… 5
5　非均匀物料设计 …… 10
5.1　非均匀物料设计的应用 …… 10
5.2　非均匀物料设计的安排 …… 11
5.3　非均匀物料试验的组织 …… 12
5.4　非均匀物料试验的统计模型 …… 12
5.5　非均匀物料试验数据的统计分析 …… 13
5.6　对数据一致性与离群值的检查 …… 15
5.7　报告非均匀物料试验的结果 …… 16
5.8　例2:非均匀物料试验 …… 16
5.9　非均匀物料设计计算的一般公式 …… 22
5.10　例3:一般公式的应用 …… 23
6　数据分析的稳健方法 …… 26
6.1　数据分析稳健方法的应用 …… 26
6.2　稳健分析:算法A …… 27
6.3　稳健分析:算法S …… 28
6.4　公式:均匀水平设计特定水平的稳健分析 …… 29
6.5　例4:均匀水平设计特定水平的稳健分析 …… 30
6.6　公式:分割水平设计特定水平的稳健分析 …… 32
6.7　例5:分割水平设计特定水平的稳健分析 …… 33
6.8　公式:非均匀物料试验特定水平的稳健分析 …… 35
6.9　例6:非均匀物料试验特定水平的稳健分析 …… 35
附录A(规范性附录)　GB/T 6379所用的符号与缩略语 …… 39
附录B(资料性附录)　算法A和算法S中所用系数的推导 …… 42
附录C(资料性附录)　稳健分析中所用公式的推导 …… 44
附录D(资料性附录)　参考文献 …… 45

前　言

GB/T 6379《测量方法与结果的准确度(正确度与精密度)》分为以下部分,其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5部分。

GB/T 6379 的本部分等同采用国际标准 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:确定标准测量方法的精密度的可替代方法》及 ISO 于 2005-06-01 发布的技术修改单 ISO 5725-5:1998/Cor. 1:2005。对 ISO 5725-5:1998 及 ISO 5725-5:1998/Cor. 1:2005 的错误作了如下的修改和更正:

——将4.8例1表6第一行中的"单元差值"更正为"单元平均值";

——将5.3.3中公式(18)中的

$$D_2 = [(\Phi^2/g) + 1/(ng)]^2/[p'(g-1)]$$

改正为:

$$D_2 = [(\gamma^2-1) + (\Phi^2/g) + 1/(ng)]^2/[p'(g-1)];$$

——将5.9中e)"样本平方和"更正为"实验室平方和";

——将6.6中公式(75)"$s_r = s^* \sqrt{2}$"更正为"$s_r = s^* /\sqrt{2}$";

——将B.1中 $c=1.5$ 时"$1/\sqrt{\theta}$"的值更正为 $c=1.5$ 时"$1/\sqrt{\beta}$"的值;

——将技术修改单 ISO 5725-5/Cor. 1:2005 中关于5.4.2中公式(25)与公式(26)中求和号上的"q"更正为"p'";

——将附录C推导公式(C.4)过程中的

$$(p - u_L - u_U) \times x^* = (p - u_L - u_U) \times x' + (u_L - u_U)s^*$$

改为:

$$(p - u_L - u_U) \times x^* = (p - u_L - u_U) \times x' + (u_L - u_U) \times 1.5s^*$$

GB/T 6379 的第1部分至第6部分作为一个整体代替 GB/T 6379—1986 及 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。

本部分的附录A为规范性附录;附录B,附录C和附录D为资料性附录。

本部分由全国统计方法应用标准化技术委员会提出并归口。

本部分起草单位:中国科学院数学与系统科学研究院、中国标准化研究院、广东出入境检验检疫局。

本部分主要起草人:冯士雍、丁文兴、姜健、于振凡、李成明、肖惠、陈玉忠。

本部分于2006年首次发布。

引　言

0.1　GB/T 6379本部分用两个术语“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的考虑,在GB/T 6379本部分中不再重复。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　GB/T 6379.2利用实验室间试验估计精密度的标准度量,即重复性标准差与再现性标准差,它给出了用均匀水平设计估计精密度的基本方法。GB/T 6379的本部分描述了可替代基本方法的其他方法。

a)　使用基本方法会产生如下风险:操作员对一个样本的测量结果会影响到后续的对同一物料的另一样本的测量结果,从而使重复性标准差和再现性标准差的估计产生偏倚。当这个风险很严重时,采用GB/T 6379本部分所描述的分割水平设计更好,因为它可减少此类风险。

b)　基本方法需要准备大量测试物料的完全相同的样本用于试验,这对非均匀物料可能是做不到的。所以使用基本方法会因样本间的变异而增大了再现性标准差的估计。GB/T 6379本部分对非均匀物料的设计由于可得到基本方法无法得到的有关样本之间变异的信息,因而在计算再现性方差的估计值时已消除了样本间的变异。

c)　基本方法要求在计算重复性标准差与再现性标准差时,对数据进行检测,剔除离群值。离群值的剔除对重复性标准差和再现性标准差的估计有时会有很大的影响。在实际中,检测离群值时,数据分析者可能不得不对哪些数据应予以剔除进行判断。GB/T 6379本部分描述了数据分析的一些稳健方法,这些方法不需对离群值进行检验并剔除,而可直接计算重复性标准差和再现性标准差。这样计算结果就不再受数据分析者判断的影响。

测量方法与结果的准确度(正确度与精密度) 第5部分:确定标准测量方法精密度的可替代方法

1 范围

GB/T 6379 的本部分:

详细描述了确定标准测量方法的重复性标准差与再现性标准差基本方法的替代方法,即分割水平设计和非均匀物料设计;

描述了用来分析精密度试验结果的稳健方法,这种方法不要求在计算过程中对数据进行离群值的检查与剔除。特别,对其中一种详尽说明了方法的使用。

GB/T 6379 的本部分是对 GB/T 6379.2 的补充,它提供在某些情况下比 GB/T 6379.2 中给出的基本方法更有价值的一些可替代的设计方法;还提供了估计重复性与再现性标准差的一种稳健分析方法,与 GB/T 6379.2 中所描述的基本方法相比,该方法依赖数据分析者的判断的程度较小。

2 规范性引用文件

下列文件中的条款通过 GB/T 6379 的本部分的引用而成为本部分的条款。凡是注日期的引用文件,其随后所有的修改单(不包括勘误的内容)或修订版本均不适用于本部分,然而,鼓励根据本部分达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件,其最新版本适用于本部分。

GB/T 3358.1—1993 统计学术语 第一部分:一般统计术语

GB/T 3358.3—1993 统计学术语 第三部分:试验设计术语

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

3 定义

GB/T 3358.1、GB/T 3358.3 与 GB/T 6379.1 中给出的定义在 GB/T 6379 的本部分中仍适用。

GB/T 6379 中使用的符号由附录 A 给出。

4 分割水平设计

4.1 分割水平设计的应用

4.1.1 GB/T 6379.2 中描述的均匀水平设计,对每个参与试验的实验室,在每个试验水平上都要求对受试物料两个或两个以上完全相同的样本进行测试。采用这种设计有如下风险:操作员在对一个样本进行测量时,测量结果可能会影响对相同物料的后续样本的测量结果。此种情形一旦发生,精密度试验结果将被歪曲:重复性标准差 σ_r 的估计值将会减小;而实验室间标准差 σ_L 的估计值将会增大。在分割水平设计中,对每一测试水平,为每个参与试验的实验室提供两种相似物料的两个样本,告诉操作员两个样本是不同的,但不告诉他们差别有多大。这样,分割水平设计提供了一种能减少前述风险的确定标

准测量方法的重复性标准差与再现性标准差的方法。

4.1.2 在分割水平试验的一个水平上所得到的数据可用来绘图，在图中，两种不同但相似物料的数据分别作为横坐标和纵坐标，图 1 即是其中一例。这样的图能够帮助识别那些相对于其他实验室偏倚最大的实验室。上述识别在有可能调查最大实验室偏倚的原因并对此采取纠正行动时，是有用的。

4.1.3 通常测量方法的重复性标准差与再现性标准差依赖于物料的水平。例如，当测量结果是由化学分析所得出的一种元素的比例时，重复性标准差与再现性标准差常随元素比例的增加而增大。对分割水平试验，在试验的同一水平上的两种物料非常相似，可以认为它们具有相同的重复性标准差和再现性标准差。对于分割水平设计而言，对试验的同一水平的两种物料得到几乎完全相同的测量结果是可以接受的。安排两种相差较大的物料反而没有意义。

在很多化学分析方法中，含有所关心成分的基体能影响精密度。所以对分割水平试验，对试验的每一个水平，都需要有两种相似基体的物质。有时能通过外加少许所关心成分物质的方法制备足够相似的物质。当物料是自然的或制造的产品时，可能难以制备分割水平试验所需的足够相似的两个产品。一种可能的解决方法是分别从同种的两批产品中抽取。应该记住，选择用于分割水平设计的物料的目的是给操作员提供不认为完全相同的样本。

4.2 分割水平设计安排

4.2.1 分割水平设计安排如表 1。

参加试验的 p 个实验室，每个在 q 个水平上均测量两个样本。

同一水平的两个样本用 a 和 b 表示，其中 a 表示一种物料的样本，而 b 表示另一种相似物料的样本。

4.2.2 分割水平试验的数据表示为 y_{ijk}，其中

下标 i 表示实验室($i=1,2,\cdots,p$)；

下标 j 表示水平($j=1,2,\cdots,q$)；

下标 k 表示样本($k=a$ 或 b)。

4.3 分割水平试验的组织

4.3.1 当计划一个分割水平试验时，应遵循 GB/T 6379.1—2004 第 6 章所给的指南。GB/T 6379.1—2004 的 6.3 中包含许多公式(公式中含有一个通常用 A 表示的量)。那些公式通常用来确定试验应包含的实验室数。分割水平试验的相应公式罗列如下。

注：这些公式由 GB/T 6379.1—2004 中注 24 中所描述的方法确定。

为评估重复性标准差与再现性标准差估计值的不确定度，计算下面诸量：

对重复性

$$A_r = 1.96\sqrt{\frac{1}{2(p-1)}} \qquad \cdots\cdots (1)$$

对再现性

$$A_R = 1.96\sqrt{\frac{[1+2(\gamma^2-1)]^2+1}{8\gamma^4(p-1)}} \qquad \cdots\cdots (2)$$

其中 $\gamma=\sigma_R/\sigma_r$。

如果 GB/T 6379.1—2004 中式(9)和式(10)中的重复次数 $n=2$，那么可以看出 GB/T 6379.1—2004 中的式(9)和式(10)与上述式(1)和式(2)相同，除了有时此处用 $p-1$ 代替 GB/T 6379.1—2004 中的 p。这是很小的差别，所以 GB/T 6379.1—2004 中的表 1、图 B.1 和 B.2 可用来评估分割水平试验的重复性和再现性标准差估计值的不确定度。

为评估分割水平试验的测量方法偏倚估计值的不确定度，计算 GB/T 6379.1—2004 定义的量 A，其中 $n=2$，(或用 GB/T 6379.1—2004 的表 2)，按 GB/T 6379.1 中描述的那样使用这个量。

为评估分割水平试验的实验室偏倚估计值的不确定度，计算 GB/T 6379.1—2004 的式(16)中定义的量 A_W，其中 $n=2$。因为在分割水平试验中重复次数为 2，所以不可能通过增加重复次数来减少实验室偏倚的估计值的不确定度。(如果必须减少该不确定度，则应用均匀水平设计来代替分割水平设计)。

4.3.2 有关分割水平试验组织的详细内容应遵照 GB/T 6379.2—2004 中第 5 章和第 6 章的指南。在 GB/T 6379.2 中，重复次数 n 取分割水平设计的分割水平数，即 2。

样本 a 与样本 b 应分别按某种随机化作业随机地分配给试验的参与者。

在分割水平试验中，统计专家在报告上来的数据中，对每个试验水平，必须能够区分哪个结果是物料 a 的，哪个是物料 b 的。因此对样本进行标记，但注意不要将这一信息透漏给试验的参与者。

表 1　分割水平试验设计的数据整理推荐格式

实验室	水平									
	1		2			j			q	
	a	b	a	b		a	b		a	b
1										
2										
i										
p										

4.4 统计模型

4.4.1 GB/T 6379 本部分使用的基本模型由 GB/T 6379.1—2004 第 5 章的式(1)给出。为估计测量方法的准确度(正确度和精密度)，假定对给定的受试物料，每个测量结果是三个分量之和：

$$y_{ijk} = m_j + B_{ij} + e_{ijk} \qquad \cdots\cdots(3)$$

式中：

m_j——给定水平 j 的总平均值(期望)，$j=1,\cdots,q$；

B_{ij}——给定实验室 i，给定水平 j 在重复性条件下偏倚的实验室分量，$i=1,\cdots,p$；$j=1,\cdots,q$；

e_{ijk}——重复性条件下，第 i 个实验室在第 j 个水平得到的第 k 个测试结果的随机误差，$k=1,\cdots,n$。

4.4.2 对分割水平试验，上述模型成为：

$$y_{ijk} = m_{jk} + B_{ij} + e_{ijk} \qquad \cdots\cdots(4)$$

与 4.4.1 中式(3)的唯一区别是，对水平 j，式(4)中的 m_{jk} 的下标 k 表明总平均值依赖于物料 a 或 b ($k=1$ 或 2)。

B_{ij} 没有下标 k 表明对一个水平，假定与 i 相应的实验室偏倚不依赖于物料 a 或 b。这就是要求两种物料相似的重要原因。

4.4.3 定义单元平均值及单元差值分别为：

$$y_{ij} = (y_{ija} + y_{ijb})/2 \qquad \cdots\cdots(5)$$

$$D_{ij} = y_{ija} - y_{ijb} \qquad \cdots\cdots(6)$$

4.4.4 分割水平试验在水平 j 的总平均值定义为：

$$m_j = (m_{ja} + m_{jb})/2 \qquad \cdots\cdots (7)$$

4.5 分割水平试验数据的统计分析

4.5.1 将数据置于表 1 所示的表中。表中每个实验室与每个水平的组合构成一个“单元”，包含 y_{ija} 和 y_{ijb} 两个数据。

计算单元差值 D_{ij} 并将它们置于表 2 所示的表中。分析方法要求按 $a-b$ 这样相同的方式计算每个差值，并保留差值的符号。

计算单元平均值 y_{ij}，并将它们置于表 3 所示的表中。

4.5.2 如果表 1 中的单元没有两个测试结果(例如由于样本损坏或数据因经后面描述的离群值检验后被剔除)，那么表 2 和表 3 中的相应的单元均应置空。

4.5.3 对每个试验水平 j，计算表 2 中第 j 列差值 D_{ij} 的平均值 D_j 与标准差 s_{Dj}：

$$D_j = \sum D_{ij}/p \qquad \cdots\cdots (8)$$

$$s_{Dj} = \sqrt{\sum (D_{ij} - D_j)^2/(p-1)} \qquad \cdots\cdots (9)$$

此处 $\sum$ 表示对实验室 $i=1,2,\cdots,p$ 求和。

若表 2 中有空白单元，p 为表 2 第 j 列中有数据的单元数，求和则对所有非空白单元进行。

4.5.4 对于每个试验水平 j，计算表 3 第 j 列平均值 y_{ij} 的平均 y_j 与标准差 s_{yj}：

$$y_j = \sum y_{ij}/p \qquad \cdots\cdots (10)$$

$$s_{yj} = \sqrt{\sum (y_{ij} - y_j)^2/(p-1)} \qquad \cdots\cdots (11)$$

此处 $\sum$ 表示对实验室 $i=1,2,\cdots,p$ 求和。

若表 3 中有空白单元，p 为表 3 中第 j 列中有数据的单元数，求和则对所有非空白单元进行。

4.5.5 利用表 2 与表 3 及按 4.5.3 与 4.5.4 计算的统计量，用 4.6 中描述的方法对数据进行一致性与离群值检验。若有数据被拒绝，则应重新计算统计量。

4.5.6 按下面的公式计算重复性标准差 s_{rj} 和再现性标准差 s_{Rj}：

$$s_{rj} = \frac{s_{Dj}}{\sqrt{2}} \qquad \cdots\cdots (12)$$

$$s_{Rj}^2 = s_{yj}^2 + s_{rj}^2/2 \qquad \cdots\cdots (13)$$

4.5.7 检查 s_{rj} 与 s_{Rj} 是否依赖于平均值 y_j。若是，使用 GB/T 6379.2—2004 中的 7.5 中描述的方法确定其函数关系。

表 2 分割水平设计单元差值列表的推荐格式

实验室	水平					
	1	2		j		q
1						
2						
i						
p						

表 3　分割水平设计单元平均值列表的推荐格式

实验室	水平					
	1	2		j		q
1						
2						
i						
p						

4.6　对数据一致性与离群值的检查

4.6.1　利用 GB/T 6379.2—2004 中的 7.3.1 中描述的 h 统计量检验数据的一致性。

检查单元差值的一致性的 h 统计量为：

$$h_{ij} = (D_{ij} - D_j)/s_{Dj} \qquad \cdots\cdots (14)$$

检查单元平均值的一致性的 h 统计量为：

$$h_{ij} = (y_{ij} - y_j)/s_{yj} \qquad \cdots\cdots (15)$$

为了揭示不一致的实验室，按实验室分组，将这些统计量按水平顺序描图，如图 2 和图 3 所示。对这些图的解释在 GB/T 6379.2—2004 的 7.3.1 中已经作了充分的讨论。如果一个实验室比其他实验室的重复性差，则在图中将会出现许多单元差大的 h 统计量。若一个实验室的测量结果普遍有偏，那么根据单元平均值得到的 h 统计量大半都在图的一个方向。上述两种情况的任一种发生，都应该要求实验室进行调查并将他们的发现反馈给试验的组织者。

4.6.2　使用 GB/T 6379.2—2004 的 7.3.4 中描述的格拉布斯检验来检验数据的歧离值和离群值。

为检验单元差中的歧离值和离群值，将表 2 的每一列作格拉布斯检验。

为检验单元平均值中的歧离值和离群值，将表 3 的每一列作格拉布斯检验。

在 GB/T 6379.2—2004 的 7.3.2 中对这些检验的解释进行了全面的讨论。这些检验用来识别与试验中报告的其他数据非常不一致的数据，如果在计算重复性和再现性标准差的过程中包含这些数据将会显著地影响这些统计量。通常，判为离群值的数据应在计算时予以剔除；而判为歧离值的数据，除非另有充分的理由则予以保留。若检验表明，表 2 和表 3 其中之一的值在计算重复性和再现性标准差时被剔除，那么在另一个表中相应的值在计算时也应予以剔除。

4.7　报告分割水平试验的结果

4.7.1　GB/T 6379.2—2004 的 7.7 提出的建议如下：

——将统计分析结果向领导小组报告；

——由领导小组作决定；

——准备一份全面报告。

4.7.2　GB/T 6379.1—2004 的 7.1 推荐了发布标准测量方法的重复性与再现性标准差的使用格式。

4.8　例 1：分割水平试验——蛋白质的测定

4.8.1　表 4 是通过氧化测定膳食中蛋白质含量的试验数据[5]。有 9 个实验室参与试验，试验有 14 个水平。每一个水平使用两份蛋白质含量组成相似的膳食。

4.8.2　表 5 和表 6 分别是根据 4.5.1 中的方法，用试验的第 14 个水平的数据计算的单元平均值和差值。

利用表 5 中的差值及 4.5.3 中的式(8)和式(9)，可得到：

$$D_{14} = 8.34\%$$

$$s_{D14} = 0.436\ 1\%$$

而将4.5.4中的式(10)和式(11)应用于表6中的单元平均值，可得到：

$$y_{14} = 85.46\%$$

$$s_{y14} = 0.453\ 4\%$$

于是利用4.5.6中的式(12)和式(13)，得重复性与再现性标准差为：

$$s_{r14} = 0.31\%$$

$$s_{R14} = 0.50\%$$

表7给出了其他水平的计算结果。

4.8.3 图1表示表4中，对水平14而言，每个实验室中的样本 a 结果对样本 b 结果的尧敦图(Youden plot)。表示实验室5的点位于图的左下角，而实验室1的点位于右上角。这表明来自实验室5的对样本 a 和样本 b 的数据有一致的负偏倚；而来自实验室1的数据对两个样本有一致的正偏倚。当用分割水平设计的数据作像图1这种点图时常可以发现这类模式。图1也表明了实验室4的结果异常，因为表示该实验室的点与表明两个样本相等的直线距离较远，其他的实验室的点均在点图的中间。图1提供了一个应调查造成3个实验室偏倚异常原因的案例。

注：关于尧敦图解释的进一步内容，参见[7]和[8]。

4.8.4 对水平14，按4.6.1中描述的方法计算的 h 统计量见表5和表6，所有水平的 h 统计量的点图见图2和图3。

在图3中，实验室5对所有水平的单元平均值的 h 统计量均为负，这说明数据一致地有负偏倚。而实验室8和9的 h 统计量几乎全为正，这说明它们的数据一致地有正偏倚(正偏倚的程度比实验室5的负偏倚小一些)。此外，图还显示实验室1、2与6有随水平而变化的实验室偏倚。实验室和水平间的这种交互作用可能提供了有关实验室偏倚原因的一些线索。

图2没有揭示出任何值得注意的模式。

4.8.5 表8给出了格拉布斯统计量的值。检验再次表明实验室5的数据可疑。

4.8.6 从分析的观点看，在对数据进行进一步分析之前，统计专家应着手对造成实验室5的可疑数据的原因进行调查。若找不到原因，就应在计算重复性和再现性标准差时剔除实验室5的所有数据。然后继续进行分析，研究重复性与再现性标准差和总平均值之间可能的函数关系，由于这些问题均已包含在GB/T 6379.2—2004中，所以此处不再考虑。

表4 例1：膳食中蛋白质含量的确定，以百分数表示

实验室	水平									
	1		2		3		4		5	
	a	b	a	b	a	b	a	b	a	b
1	11.11	10.34	10.91	9.81	13.74	13.48	13.79	13.00	15.89	15.26
2	11.12	9.94	11.38	10.31	14.00	13.12	13.44	13.06	15.69	15.10
3	11.26	10.46	10.95	10.51	13.38	12.70	13.54	13.18	15.83	15.73
4	11.07	10.41	11.66	9.95	13.01	13.16	13.58	12.88	15.08	15.63
5	10.69	10.31	10.98	10.13	13.24	13.33	13.32	12.59	15.02	14.90
6	11.73	11.01	12.31	10.92	14.01	13.66	14.04	13.64	16.43	15.94
7	11.13	10.36	11.38	10.44	12.94	12.44	13.63	13.06	15.75	15.56
8	11.21	10.51	11.32	10.84	13.09	13.76	13.85	13.49	15.98	15.89
9	11.80	11.21	11.35	9.88	13.85	14.46	13.96	13.77	16.51	15.72

表 4(续)

实验室	水平									
	6		7		8		9		10	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
1	20.14	19.78	20.33	20.06	46.45	44.42	52.05	49.40	65.84	59.14
2	19.25	20.25	20.36	19.94	46.69	44.62	51.94	48.81	66.31	59.19
3	20.48	19.86	20.56	20.11	46.90	44.56	52.18	48.90	66.06	58.52
4	21.54	20.06	20.64	20.46	47.13	46.29	51.73	48.56	65.93	58.93
5	19.90	19.66	20.56	19.24	45.83	43.73	50.84	47.91	64.19	57.94
6	20.31	20.27	20.85	20.63	46.86	43.96	52.18	49.03	65.73	58.77
7	20.00	20.56	20.25	20.19	46.25	44.31	52.25	49.44	66.06	59.19
8	20.43	20.69	20.85	20.27	47.11	44.40	52.44	48.81	65.66	59.38
9	20.64	21.01	20.78	20.89	47.09	45.15	52.19	48.46	66.33	59.47

实验室	水平									
	11		12		13		14			
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*		
1	84.16	80.86	85.38	81.71	87.64	88.23	90.24	82.10		
2	84.50	81.06	85.56	82.44	88.81	88.38	89.88	81.44		
3	82.26	79.43	85.26	82.15	88.58	88.12	89.48	81.67		
4	84.39	80.08	85.20	81.76	88.47	87.98	90.04	80.73		
5	81.71	79.01	83.58	79.74	86.43	86.19	88.59	80.46		
6	82.85	81.16	84.44	80.90	87.78	86.89	89.40	80.88		
7	86.25	81.00	84.88	81.44	88.06	88.00	89.31	81.38		
8	84.59	81.16	84.96	81.71	88.50	87.98	89.94	81.56		
9	83.05	80.93	84.73	81.94	88.24	88.05	89.75	81.35		

表 5　例 1：水平 14 的单元差值

实验室	单元差值 %	*h* 统计量
1	8.14	−0.459
2	8.44	0.229
3	7.81	−1.215
4	9.31	2.224
5	8.13	−0.482
6	8.52	0.413
7	7.93	−0.940
8	8.38	0.092
9	8.40	0.138

表 6　例 1：水平 14 的单元平均值

实验室	单元平均值 %	*h* 统计量
1	86.170	1.576
2	85.660	0.451
3	85.575	0.263
4	85.385	−0.156
5	84.525	−2.052
6	85.140	−0.696
7	85.345	−0.244
8	85.750	0.649
9	85.550	0.208

表 7　例 1:根据表 4 中所有 14 个水平数据计算的平均值、平均差值及标准差

水平 j	实验室数 p	总平均值 y_j %	平均差值 D_j %	标准差 s_{rj} %	s_{Dj} %	s_{Lj} %	s_{Rj} %
1	9	10.87	0.73	0.35	0.21	0.15	0.36
2	9	10.84	1.05	0.36	0.43	0.30	0.42
3	9	13.41	0.13	0.44	0.55	0.39	0.52
4	9	13.43	0.50	0.30	0.21	0.15	0.32
5	9	15.66	0.27	0.39	0.40	0.29	0.44
6	9	20.27	0.06	0.40	0.73	0.52	0.54
7	9	20.39	0.38	0.30	0.41	0.29	0.37
8	9	45.60	2.21	0.44	0.37	0.26	0.47
9	9	50.40	3.16	0.44	0.35	0.25	0.47
10	9	62.37	6.84	0.53	0.40	0.28	0.57
11	9	82.14	3.23	1.01	1.08	0.77	1.15
12	9	83.17	3.45	0.74	0.46	0.33	0.77
13	9	87.91	0.30	0.69	0.41	0.29	0.72
14	9	85.46	8.34	0.45	0.44	0.31	0.50

表 8　例 1:格拉布斯统计量

差值的格拉布斯统计量				
水平	一个最小值	两个最小值	两个最大值	一个最大值
1	1.653	0.508 1	0.313 9	2.125
2	1.418	0.394 5	0.473 8	1.535
3	1.462	0.362 8	0.532 3	1.379
4	1.490	0.584 1	0.477 1	1.414
5	2.033	0.348 5	0.607 5	1.289
6	1.456	0.549 0	0.321 0	1.947
7	1.185	0.682 0	0.171 2	2.296*(5)
8	0.996	0.757 1	0.141 8*(6;8)	1.876
9	1.458	0.500 2	0.309 2	1.602
10	1.474	0.336 0	0.457 8	1.737
11	1.422	0.508 9	0.294 3	1.865
12	1.418	0.600 9	0.289 9	1.956
13	2.172	0.232 5	0.632 6	1.444
14	1.215	0.622 0	0.236 2	2.224*(4)
单元平均值的格拉布斯统计量				
水平	一个最小值	两个最小值	两个最大值	一个最大值
1	1.070	0.660 7	0.129 1*(6;9)	1.832
2	1.318	0.628 8	0.211 8	2.165
3	1.621	0.477 1	0.407 7	1.680

表 8(续)

单元平均值的格拉布斯统计量				
水平	一个最小值	两个最小值	两个最大值	一个最大值
4	1.591	0.533 9	0.380 7	1.429
5	1.794	0.401 8	0.500 9	1.333
6	1.291	0.494 7	0.409 5	1.386
7	1.599	0.503 6	0.439 1	1.470
8	1.872	0.375 3	0.453 6	1.404
9	2.328*(5)	0.131 7*(4;5)	0.741 7	1.025
10	2.456**(5)	—	—	1.000
11	1.756	0.246 9	0.575 9	1.472
12	2.037	0.106 3*(5;6)	0.711 6	1.130
13	2.308*(5)	0.073 3**(5;6)	0.777 7	0.994
14	2.052	0.278 1	0.548 6	1.576

注：括号内的数是指产生了歧离值或离群值的实验室号。

对于 9 个实验室情形，格拉布斯统计量(无论对于差值还是单元平均值)的临界值如下：

	歧离值 (*)	离群值 (**)
对一个离群值的格拉布斯检验	2.215	2.387
对两个离群值的格拉布斯检验	0.149 2	0.085 1

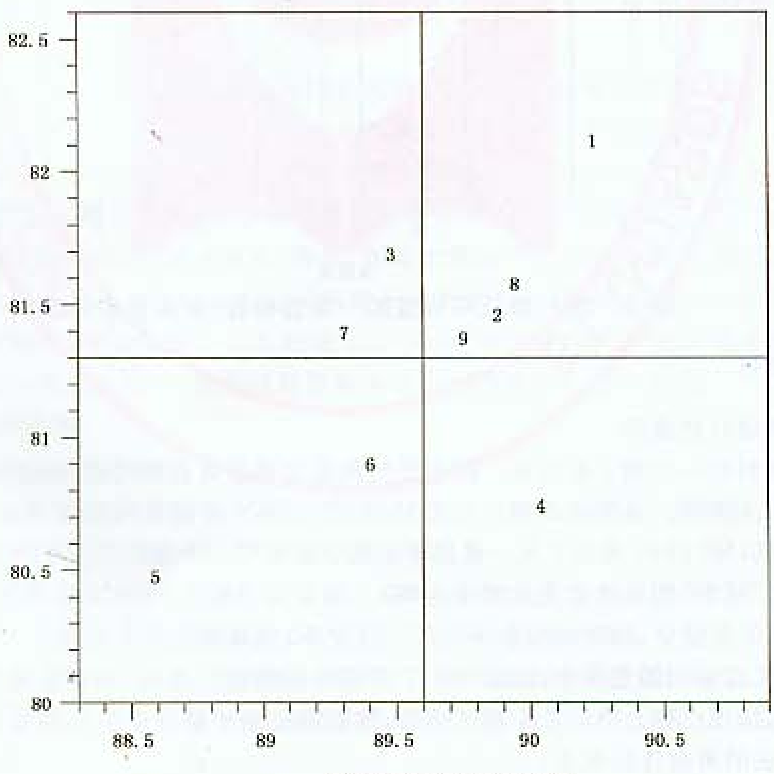


图 1　例 1:水平 14 的数据

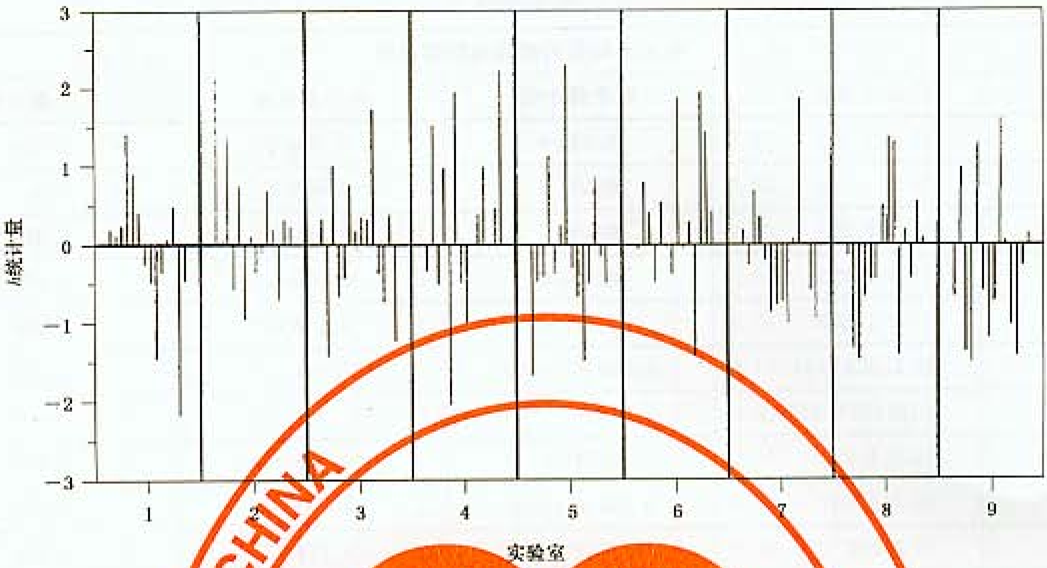


图 2 例 1:单元差值的一致性检验(以实验室分组)

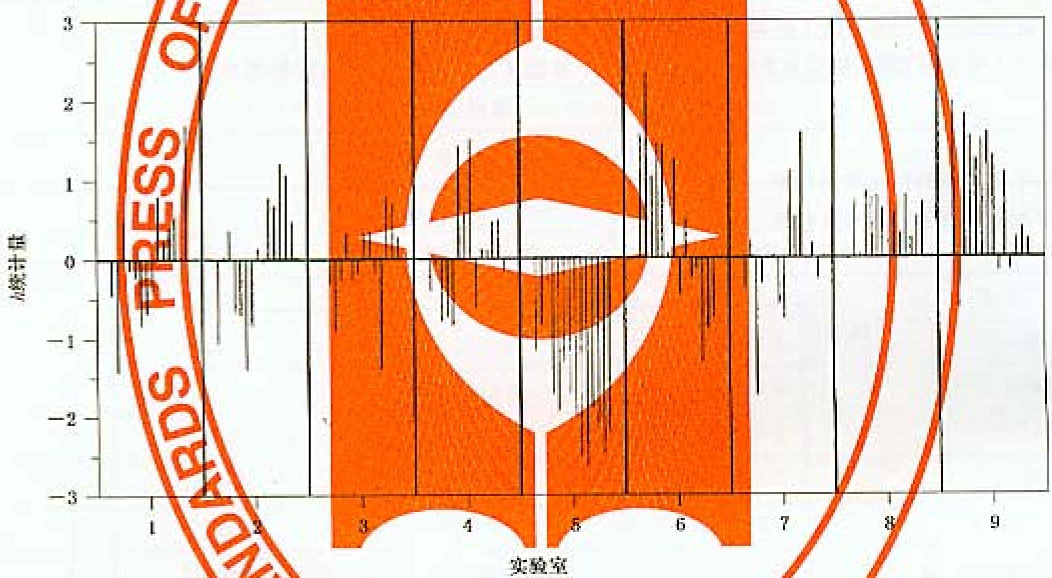


图 3 例 1:单元平均值的一致性检验(以实验室分组)

5 非均匀物料设计

5.1 非均匀物料设计的应用

5.1.1 非均匀物料的一个例子是皮革。没有任何两张皮革是完全相同的,即使是一张皮革中的不同部位,其性质也有很大差异。通常皮革使用的检验是 BS 3144[3] 中的抗张强度测试。测试是对一个哑铃型的试样上进行的(BS 3144 规定了从一张皮革上切下的试样的个数以及它们在皮革中的位置和定向,所以皮革测试中,“样本”的自然定义是整张皮革)。如果用 GB/T 6379.2 中所述的均匀水平设计进行精密度试验,给每个实验室、每个试验水平分配一张皮革,从每张皮革上得到 2 个测试结果,那么皮革间的变异就会加进实验室间的变异中,因而增大了再现性标准差。然而,如果给每个实验室、每个试验水平分配两张皮革,每张皮革上得到 2 个测试结果,则数据能用于估计皮革间的变异,计算得到去除皮革间变异的测量方法的再现性标准差。

5.1.2 非均匀物料的另一个例子是沙子(它可用于制造水泥)。在堆放过程中,由于空气和水的作用,处于上下不同层的沙子的粒度是不同的。所以在使用沙子时,人们总关心沙粒大小的分布。在水泥制

造技术中，沙粒大小分布是用网格测试来度量的(例如 BS 812-103[1])。为此，先抽取一份样，从份样中产生一个或多个试样。典型的份样质量约为 10 kg，试样质量约为 200 g。由于物料的自然变异，在同一产品的不同份样间会有差异。因此，如同皮革一样，若进行均匀水平试验，即给每个实验室在每个水平下分配一个份样，则份样之间的变异将会使得测量方法的再现性标准差加大。但如果在每个水平给每个实验室分配 2 个份样，则可计算去除份样间变异的再现性标准差。

5.1.3 上面两个例子也显示了非均匀物料的另一个性质：由于物料的变化，试样或试样部分准备可能是一个重要的变异源。对皮革而言，从一张皮革上切割试样的过程就会对测量抗张强度产生很大的影响。利用沙子的网格检验，从份样中准备试样部分的过程在该检验方法中通常是主要的变异源。如果按与通常实际情况不一样的方式(即试图产生相同的“样本”)将试样或试样部分准备用于精密度试验，那么由这样的试验得出的重复性和再现性标准差的值将不会代表通常意义下的变异。有些情况可以通过某些特殊的过程，尽可能的消除物料的变异(例如对一个能力验证，或是在开发一个测量方法过程中，精密度试验只用作工作程序的一部分)，从而有可能产生相同的“样本”。然而，当精密度试验的目的是发现变异，而该变异可在实际中发生时(例如，当买卖双方检验相同产品的样本时)，则物料的非匀质性所产生的变异必然包含在测量方法精密度中。

为确保试验中的每一测试结果与其他测试结果相互独立，应谨慎行事。如果几个试样分别是在不同的试样准备阶段上准备的，则不能保证测试结果之间的独立性，所以由试样准备所产生的偏倚或偏离将会极大影响这些试样的测试结果。

5.1.4 本章提出的非均匀物料的设计能得到有关样本间变异的信息，而这些信息在 GB/T 6379.2—2004 中描述的均匀水平设计却得不到。获得额外的信息必然会发生费用问题：因为非均匀物料的设计需要检验更多的样本。这些额外的信息可能很有价值。在 5.1.1 中讨论的皮革例子中，关于皮革之间变异的信息可能用来决定在评估交付物的质量时应该使用多少张皮革，即决定是在每张皮革上制作较少的试样，从而使用较多张的皮革还是在每张皮革上制作较多的试样，从而使用较少张的皮革。在 5.1.2 中讨论的沙子的例子中，关于份样间变异的信息可用来决定抽取份样的过程是令人满意的还是需要进行改进。

5.1.5 本章所描述的设计用于按层次排列的三因素试验：第一个因素是层次最高的“实验室”，第二个因素是下一层次的“实验室内样本”，最后一个因素是层次中最低的“样本测试结果”。实际中可能遇到的另一种情形是按如下层次排列的三因素：层次最高的“实验室”，中间层次的“实验室内的测试结果”，以及最低层次的“测试结果的测定”。例如给每个参与精密度试验的实验室分发一个均匀物料的样本，要求每个实验室对样本上进行两次(或更多次)测试，而每次测试又包括多次测定，测试结果以这些测定的平均值表示。在 5.5、5.6 及 5.9 中给出的公式可用于这类试验所获得的数据，但是重复性与再现性标准差计算公式与此处稍有不同(参见 5.5.5 的注 2)。另外，对为获得一个测试结果需进行多次测定以进行平均的测定数也必须给以规定，因为这将影响重复性和再现性标准差的值。

5.2 非均匀物料设计的安排

5.2.1 非均匀物料的设计的安排如表 9 所示。

对 p 个实验室，q 个水平，给每个参加试验的实验室在每个水平上提供 2 个样本。这样，每个试验单元包含 4 个测试结果(2 个样本每个都有 2 个测试结果)。

通过考虑每个实验室在每个水平上分配 2 个以上样本，或每个样本上进行两次以上测量，有可能将这个简单设计进行推广。对更一般的设计，需要的计算要比每个样本只有 2 个测试结果及每个实验室每个水平有 2 个样本的计算更为复杂。然而更一般设计与简单设计的原理是相同的。所以此处计算将仅仅对简单设计进行详细讨论。对一般设计，重复性与再现性标准差值的计算公式在 5.9 中给出，在 5.10 中给出其应用实例。

5.2.2 非均匀物料设计的数据表示为 y_{ijtk}，其中：

下标 i 表示实验室($i=1,2,\cdots,p'$)；

下标 j 表示水平($j=1,2,\cdots,q$);

下标 t 表示样本($t=1,2,\cdots,g$);

下标 k 表示试验结果($k=1,2,\cdots,n$)。

通常,$g=2$,$n=2$。在更一般的设计中,$g\geqslant 2$,$n\geqslant 2$。

注:在 GB/T 6379.1 和 GB/T 6379.2 中,p 既被用作实验室数,也用作柯克伦检验的临界值表的检索数;对均匀水平试验,这两个数相同。对非均匀物料设计,柯克伦检验的检索数可为实验室数的倍数,所以此处用 p' 表示实验室数,而 p 仍表示为柯克伦检验的检索数。

5.3 非均匀物料试验的组织

5.3.1 当计划一个非均匀物料试验时,应遵循 GB/T 6379.1—2004 第 6 章的指南,需考虑的另一个问题是:

对每个实验室每个水平应准备多少个样本?

由于考虑到费用,样本数通常为 2。

GB/T 6379.1—2004 第 6 章和附录 B 中的公式及图表可用来帮助选择实验室数、样本数和重复次数,但须按 5.3.2 至 5.3.5 进行修正。

5.3.2 由非均匀物料试验得到的重复性标准差估计值的不确定度,可以根据计算 GB/T 6379.1—2004 的 6.3 中引入的量 A_r 进行评定:

$$A_r = 1.96\sqrt{1/[2p'g(n-1)]} \qquad \cdots\cdots(16)$$

它替代 GB/T 6379.1—2004 中式(9)所定义的 A_r。然而,式(16)可由 GB/T 6379.1—2004 的式(9)用 $p'\times g$ 代替 p 得出。因此当用 GB/T 6379.1—2004 中图 B.1 和表 1,查重复性标准差估计值所需的 A_r 时,可以 $p'\times g$ 代替那里的 p。在给每个实验室的每个水平制备 $g=2$ 个样本的通常情形,以 $p=2p'$ 查询 GB/T 6379.1 中的图和表。

注:上述的 A_r(或下述的 A_R)的公式均根据 GB/T 6379.1—2004 的注 24 的方法得出。

5.3.3 由非均匀物料试验得到的再现性标准差估计值的不确定度,可以根据计算 GB/T 6379.1—2004 的 6.3 中引入的量 A_R 进行评定:

$$A_R = 1.96\sqrt{(D_1+D_2+D_3)/(2\gamma^4)} \qquad \cdots\cdots(17)$$

它替代 GB/T 6379.1—2004 的式(10)所定义的 A_R。其中

$$D_1 = [(\gamma^2-1)+(\Phi^2/g)+1/(ng)]^2/(p'-1)$$

$$D_2 = [(\gamma^2-1)+(\Phi^2/g)+1/(ng)]^2/[p'(g-1)]$$

$$D_3 = 1/[p'(g-1)]$$

$$\Phi = \sigma_H/\sigma_r \qquad (\sigma_H \text{ 将在后面的 5.4.1 中定义})$$

$$\gamma = \sigma_R/\sigma_r \qquad \cdots\cdots(18)$$

Φ 和 γ 的值可根据标准差 σ_H 的初始估计值来确定。σ_R 和 σ_r 可在标准化测量方法过程中获得。

5.3.4 非均匀物料试验组织的详细内容应遵循 GB/T 6379.2—2004 第 5 章和第 6 章所给的指南。

GB/T 6379.2—2004 的 5.1.2 包括了对"n 个测试组"或"n 个测量组"的要求(例如 n 个测试的组应在重复性条件下进行的)。在一个非均匀物料的试验中,这些要求涉及在一个单元内的 $g\times n$ 测试组中,即设计在一个实验室中一个水平的所有测试。

在一个非均匀物料的试验中,对每个水平必须制备的样本数是 $p'\times g$(即在通常情况下,当 $g=2$ 时为 $2p'$)。重要的是将 $p'\times g$ 个样本随机地分配给参加试验的实验室。

5.4 非均匀物料试验的统计模型

5.4.1 GB/T 6379 本部分所用的基本模型是 4.4.1 中式(3)。对非均匀物料试验,这个模型扩展为:

$$y_{ijtk} = m_j + B_{ij} + H_{ijt} + e_{ijtk} \qquad \cdots\cdots(19)$$

分量 m、B 和 e 与 4.4.1 中式(3)的意义相同,但式(19)包含了一个额外的分量 H_{ijt},它表示样本间的变异,下标 t 表示实验室内样本(其他下标的含义已在 5.2.2 中给出)。

有理由假定样本间的变异是随机的，它不依赖实验室，但可能依赖于试验的水平，所以分量 H_{ijt} 的期望为 0，方差为：

$$\mathrm{var}(H_{ijt}) = \sigma_{Hj}^2 \qquad \cdots\cdots(20)$$

5.4.2 在每个实验室有 2 个样本，每个样本有 2 个测试结果($g=n=2$)的通常情形，令：

a) 实验室 i，水平 j 和样本 t ($t=1$ 或 2)的样本平均值及测试结果间极差为：

$$y_{ijt} = (y_{ijt1} + y_{ijt2})/2 \qquad \cdots\cdots(21)$$

$$w_{ijt} = |\, y_{ijt1} - y_{ijt2} \,| \qquad \cdots\cdots(22)$$

b) 实验室 i，水平 j 的单元平均值及样本间极差为：

$$y_{ij} = (y_{ij1} + y_{ij2})/2 \qquad \cdots\cdots(23)$$

$$w_{ij} = |\, y_{ij1} - y_{ij2} \,| \qquad \cdots\cdots(24)$$

c) 水平 j 的总平均值及单元平均值的标准差为：

$$y_j = \sum_{i=1}^{p'} y_{ij} / p' \qquad \cdots\cdots(25)$$

$$s_{yj} = \sqrt{\sum_{i=1}^{p'} (y_{ij} - y_j)^2 / (p'-1)} \qquad \cdots\cdots(26)$$

其中求和是对所有实验室，$i=1,2,\cdots,p'$。

5.5 非均匀物料试验数据的统计分析

5.5.1 这里将详细讨论给每个实验室在每个水平上分配 2 个样本，每个样本进行两次测量这种通常情形。(一般情形将在 5.9 和 5.10 中考虑。)

将数据置入表 9 所示的表中。每个实验室和水平的每个组合构成了表中的一个“单元”，它包括 4 个测试结果。

根据 5.4.2 中的式(21)与式(26)：

a) 计算测试结果间极差，将其置入表 10 所示的表中；

b) 计算样本间极差，将其置入表 11 所示的表中；

c) 计算单元平均值，将其置入表 12 所示的表中。

所有极差都记录为正值(即忽略其符号)。

表 9 非均匀物料设计数据整理的推荐格式

实验室	样本	水平 1	水平 2		水平 j		水平 q
		测试结果					
		1 2	1 2		1 2		1 2
1	1 2						
2	1 2						
i	1 2						
p'							

表 10 非均匀物料设计测试结果间极差列表的推荐格式

实验室	样本	水平 1	水平 2		水平 j		水平 q
1	1						
	2						
2	1						
	2						
i	1						
	2						
p'							

表 11 非均匀物料设计样本间极差列表的推荐格式

实验室	水平 1	水平 2		水平 j		水平 q
1						
2						
i						
p'						

表 12 非均匀物料设计单元平均值列表的推荐格式

实验室	水平 1	水平 2		水平 j		水平 q
1						
2						
i						
p'						

5.5.2 若表 9 中的某个单元包含的测试结果数少于 4 个（例如，由于样本受损或数据根据离群值检验的结果被剔除），则

a) 根据后面给出的一般情形的公式；或

b) 忽略该单元的所有数据。

通常倾向于选用 a)，因为 b) 浪费了数据，但所用的公式简单。

5.5.3 对试验的每一个水平，计算以下各项：

a) 表 10 中第 j 列测试结果间极差的平方和(对 p' 个试验室和 2 个样本求和)：

$$SS_{rj} = \sum_{i=1}^{p'} \sum_{t=1}^{2} w_{ijt}^2 \qquad \cdots\cdots (27)$$

b) 表 11 中第 j 列的样本间极差的平方和(对 p' 个试验室求和)：

$$SS_{Hj} = \sum_{i=1}^{p'} w_{ij}^2 \qquad \cdots\cdots (28)$$

c) 表 12 中第 j 列的单元平均值的平均和标准差(利用 5.4.2 中(25)式和(26)式)。

5.5.4 根据表 10，表 11 和表 12 及 5.5.3 中计算的统计量，按 5.6 中描述的方法检验数据的一致性及离群值。若有任何数据被剔除，重新计算统计量。

5.5.5 根据以下公式计算重复性标准差 s_{rj} 及再现性标准差 s_{Rj}。

$$s_{rj}^2 = SS_{rj}/(4p') \qquad \cdots\cdots (29)$$

$$s_{Rj}^2 = s_{yj}^2 + (SS_{rj} - SS_{Hj})/(4p') \qquad \cdots\cdots (30)$$

如果计算得到的

$$s_{Rj} < s_{rj} \qquad \cdots\cdots (31)$$

则令

$$s_{Rj} = s_{rj} \qquad \cdots\cdots (32)$$

按以下公式计算度量样本间变异的标准差的估计值 s_{Hj}：

$$s_{Hj}^2 = SS_{Hj}/(2p') - SS_{rj}/(8p') \qquad \cdots\cdots (33)$$

注 1：人们常对用显著性检验考察样本间的差异在统计上是否显著感兴趣，然而这并不是分析工作必不可少的一部分。利用这样的检验决定是否可以在分析中忽略样本间的差异(以将每一单元中的测试结果都作为是在相同样本而得处理)是不正确的。这将会把偏倚引入重复性标准差的估计中，因为样本间的差异在统计上不显著并不表明样本间的变异可以忽略。

注 2：在 5.1.5 中所描述的情形(当 3 个因素分别是"实验室"、"实验室内测试结果"和"测试结果的测定"时)，重复性和再现性标准差应计算如下：

$$s_{rj}^2 = SS_{Hj}/(2p')$$

$$s_{Rj}^2 = s_{yj} + SS_{Hj}/(4p')$$

这些公式适用于测试结果为 2 个测定平均值的情形。

5.5.6 检查 s_{rj} 和 s_{Rj} 是否依赖于总平均值 y_j。若是，则根据 GB/T 6379.2—2004 的 7.5 中描述的方法确定其函数关系。

5.6 对数据一致性与离群值的检查

5.6.1 根据 GB/T 6379.2—2004 的 7.3.1 中描述的方法，用 h 统计量和 k 统计量来检查数据的一致性。

为检查单元平均值的一致性，计算以下 h 统计量：

$$h_{ij} = (y_{ij} - y_j)/s_{yj} \qquad \cdots\cdots (34)$$

将这些统计量以实验室分组，按水平顺序点图，即可揭示与其他不一致的实验室。

为检查样本间极差的一致性，计算以下 k 统计量：

$$k_{ij} = w_{ij}/\sqrt{SS_{Hj}/p'} \qquad \cdots\cdots (35)$$

将这些统计量以实验室分组，按水平顺序点图，即可揭示与其他不一致的实验室。

为检查测试结果间极差的一致性，计算以下 k 统计量：

$$k_{ijt} = w_{ijt}/\sqrt{SS_{rj}/(2p')} \qquad \cdots\cdots (36)$$

将这些统计量以实验室分组，按水平顺序点图，即可揭示与其他不一致的实验室。

这些图的解释在 GB/T 6379.2—2004 的 7.3.1 中进行了充分的讨论。若一个实验室报告的结果普遍有偏，则该实验室单元平均值的 h 统计量图中，大多数统计量的数值大且方向相同。若一个实验室

的测量不是在重复性条件下的各水平上进行的(受外部因素的影响,从而增加样本间的变异),则可看到它的样本间极差 k 统计量就会异常的大;若实验室的重复性很差,则测试结果间极差的 k 统计量就异常的大。

5.6.2 用 GB/T 6379.2—2004 的 7.3.3 和 7.3.4 中描述的柯克伦(Cochran)检验和格拉布斯(Grubbs)检验来检测数据中的歧离值和(统计)离群值。

为检验测试结果间极差的歧离值和离群值,对每个水平 j,计算柯克伦统计量如下:

$$C = w_{\max}^2 / SS_{rj} \qquad \cdots\cdots(37)$$

其中 $w_{\max}$ 是水平 j 的测试结果间极差 w_{ijt} 中的最大值。

当查 GB/T 6379.2—2004 中 8.1 中的临界值表时,临界值即为对应于表左侧 $p=2p'$ 的行与顶部 $n=2$ 的列的数值。

为检验样本间极差的歧离值和离群值,对每个水平 j 计算计算柯克伦统计量如下:

$$C = w_{\max}^2 / SS_{Hj} \qquad \cdots\cdots(38)$$

其中 $w_{\max}$ 是水平 j 的样本间极差 w_{ij} 中的最大值。

当查 GB/T 6379.2—2004 中 8.1 中的临界值表时,临界值即为对应于表左侧 $p=p'$ 的行与顶部 $n=2$ 的列的数值。

为检验单元平均值的歧离值和离群值,对每个水平 j,计算由 GB/T 6379.2—2004 中 7.34 所示的由单元平均值计算的格拉布斯统计量(其中 GB/T 6379.2—2004 中的 s 是 5.4.2 中(26)式定义的 s_{yj})。

这些检验的解释在 GB/T 6379.2—2004 中的 7.3.2 中已作了说明。非均匀物料试验中,这些结果按下面顺序进行检验。第一步,将柯克伦检验应用于测试结果间的极差。若基于此检验,判定某个测试结果间极差为离群值,应予以剔除,则在计算重复性和再现性标准差时须将产生这个离群极差的 2 个测试结果都予以剔除(但是该单元中的其他测试结果仍应保留)。第二步,将柯克伦检验用于样本间极差,最后将格拉布斯检验用于单元平均值。若判定某个样本间极差为离群值,或某个单元平均值为离群值,而产生离群值的相应测试结果予以剔除,则在计算重复性和再现性标准差时,应剔除相应单元的所有测试结果。

5.7 报告非均匀物料试验的结果

4.7 的内容完全适用于非均匀物料试验。

5.8 例 2:非均匀物料试验

5.8.1 用于铺设机场和公路表面的混合材料(基于水泥质的或沥青质的)必须有良好的防水和防冻性能。度量这种性能的一种方法是进行用硫酸镁坚固性试验[2]。试验是将混合材料的试样浸泡在饱和的硫酸镁溶液中,然后进行干燥,多次重复此过程。试验开始时,制备的全部试样都不能通过 10.0 mm 的筛子。经过试验的处理后,试样颗粒的尺寸减小,测试结果以试样在试验结束时穿过 10.0 mm 筛子的质量百分数表示。结果愈高(超过 10%到 20%的颗粒穿过筛子),材料的坚固性愈差。

5.8.2 将此种混合材料的 8 个样本成对的送往 11 个实验室,在每个样本通过硫酸镁坚固性试验,得到 2 个测试结果,表 13 为所得的测试结果数据。这些样本的总质量大约为 100 kg(它们还用于许多其他的试验)。试样的质量约为 350 g。

5.8.3 表 14、表 15 和表 16 分别列出根据 5.4.2 中式(21)至式(24)计算的,第 6 个试验水平的测试结果间极差、样本间极差和单元平均值(单位:%)。

根据 5.5.3 中式(27)和式(28),从表 14 中测试结果间极差和表 15 中样本间极差,得:

$$SS_{r6} = 381.66 \qquad SS_{H6} = 160.530\,0$$

将 5.4.2 中式(25)和式(26)用于单元平均值,得:

$$y_6 = 19.0\text{(总平均值)}$$

$$s_{y6} = 5.03$$

所以,根据 5.5.5 中式(29)至式(33),可得重复性标准差、再现性标准差以及用于度量样本间变异

的标准差分别为：

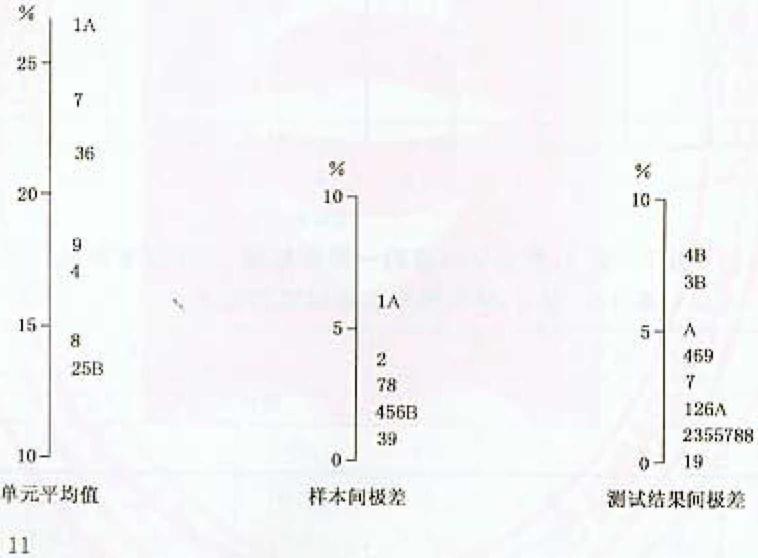
$$s_{r6}=2.95 \qquad s_{R6}=5.51 \qquad s_{H6}=1.72$$

表 17 给出了其他水平的计算结果。

5.8.4 图 4 为水平 6 的测试结果间极差、样本间极差及单元平均值的条形图。这种图对由不同来源(测试结果间、样本间和实验室间)产生的变异的大小易于理解。图 4 表明在本试验中，对水平 6，单元平均值间有的变异较大，因此若规范中列入了测试方法，由于测试结果不同，很可能在买卖双方之间产生争议。由于样本间的极差比测试结果间的极差较小，因此在水平 6，样本间的变异不重要。

5.8.5 用 5.6.1 的方法对水平 6 计算所得的统计量 h 和 k 统计量的值也分别列于表 14、15 和 16 中。所有水平的这些值则绘于图 5 至图 7 中(在图中，已对水平重新排列，如表 7，按总平均值递增顺序排列)。图 5 表明实验室 6 的测试结果间极差有几个 k 统计量值较高，说明该实验室的重复性比其他实验室差。图 6 表明有三个实验室(1，6 和 10)的样本间极差的 k 统计量值较高；这可能是因为这三个实验室并非严格地按推荐的程序从份样中制备试样。图 7 表明大多数实验室有一致正或负的 h 统计量(实验室 1、6 和 10 再次表明其值最大)。这是大多数实验室有一致偏倚的明显证据，表明测量方法没有很好规范。

5.8.6 按 5.6.2 方法，对数据进行柯克伦检验和格拉布斯检验，得到表 18 所示的结果，共检测出 2 个离群值。在缺少其他信息的情况下，应剔除这些数据，然后重新进行计算。接着可进一步按 GB/T 6379.2—2004 中对均匀设计相同的方法来研究函数关系。



A，B=实验室 10 和 11

图 4 例 2：根据表 14、15 和 16 在水平 6 的极差和平均值的条形图

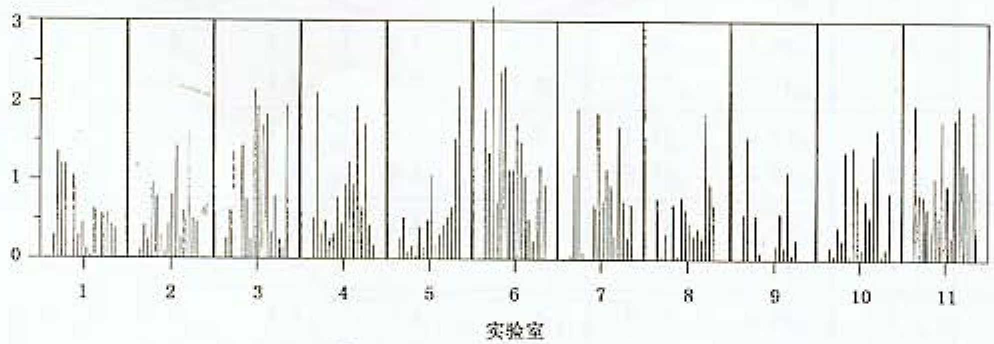


图 5 例 2：测试结果间极差的一致性检验(以实验室分组)

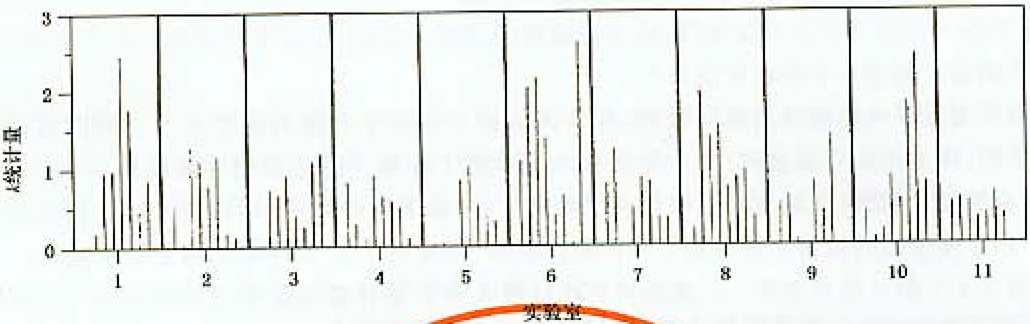


图 6　例 2:样本间极差的一致性检验(以实验室分组)

图 7　例 2:单元平均值的一致性检验(以实验室分组)

表 13　例 2:硫酸镁坚固性试验的测试结果(%)

实验室	样本	水平 1		水平 2		水平 3		水平 4	
		测试结果							
		1	2	1	2	1	2	1	2
1	1	69.2	67.0	7.4	8.0	4.1	3.5	10.4	10.1
	2	69.7	71.7	6.6	5.7	10.5	13.1	13.9	13.8
2	1	66.5	64.1	1.9	2.1	3.0	3.2	8.7	6.7
	2	65.7	65.8	4.2	3.3	1.9	1.1	8.3	4.8
3	1	68.7	69.5	6.3	5.8	2.4	2.9	11.7	7.0
	2	67.7	77.7	9.7	5.3	2.1	3.3	7.9	12.0
4	1	77.5	75.3	2.0	3.6	2.4	1.4	9.4	7.1
	2	76.3	77.2	4.7	3.8	6.4	2.3	10.7	7.7
5	1	55.4	63.2	3.8	4.1	1.3	0.8	3.7	6.3
	2	65.9	54.7	2.1	3.1	0.7	1.7	3.3	3.7
6	1	64.8	70.9	8.4	6.1	6.0	9.7	16.5	12.3
	2	78.2	73.4	8.3	10.6	12.4	9.8	13.2	16.8

表 13(续)

实验室	样本	水平 1		水平 2		水平 3		水平 4	
		测试结果							
		1	2	1	2	1	2	1	2
7	1	64.8	63.4	4.3	5.7	2.9	3.0	7.5	9.3
	2	67.0	63.4	7.7	3.9	4.3	6.4	11.1	8.3
8	1	64.9	68.4	4.4	2.8	1.3	2.8	5.7	6.8
	2	65.4	65.5	5.4	6.7	2.7	2.8	4.8	5.5
9	1	—	—	—	—	1.1	0.0	6.6	7.0
	2	—	—	—	—	0.7	3.7	4.9	6.3
10	1	57.0	57.7	3.3	0.4	2.1	2.4	5.5	5.8
	2	57.1	52.7	4.2	2.3	3.6	3.5	3.9	5.7
11	1	70.6	75.2	5.3	6.4	5.7	1.9	9.5	7.2
	2	77.9	68.2	3.5	7.1	1.4	3.0	8.1	7.4

实验室	样本	水平 5		水平 6		水平 7		水平 8	
		测试结果							
		1	2	1	2	1	2	1	2
1	1	8.9	7.4	31.1	28.5	38.7	41.7	[illegible]	4.1
	2	7.6	9.1	23.0	23.1	44.2	41.1	7.3	4.4
2	1	3.2	3.5	16.5	15.4	36.6	45.2	3.2	5.4
	2	2.8	4.0	10.3	12.8	43.2	40.5	1.7	2.5
3	1	4.4	6.1	24.3	16.7	38.9	43.1	3.7	7.7
	2	6.0	6.0	20.8	22.2	46.1	47.4	3.5	5.6
4	1	2.7	3.1	20.2	16.2	32.0	35.5	2.9	2.2
	2	2.3	2.9	20.0	11.9	26.5	35.7	3.2	2.3
5	1	[illegible]	1.4	13.8	15.1	36.7	39.5	1.1	1.2
	2	1.6	1.3	11.5	13.3	37.6	34.1	0.6	1.7
6	1	8.2	4.2	20.3	24.7	49.4	50.6	11.9	18.5
	2	3.7	4.6	21.0	18.9	48.2	52.4	14.9	8.1
7	1	3.1	5.5	27.2	23.3	38.9	29.9	—	1.7
	2	5.6	5.5	21.5	22.7	34.4	38.3	2.2	5.0
8	1	1.8	2.2	13.6	12.0	27.0	37.0	0.3	2.2
	2	4.0	4.0	15.6	16.7	39.7	34.6	3.6	3.7
9	1	3.8	3.8	17.7	17.1	33.4	33.1	1.8	2.0
	2	3.5	2.8	21.4	16.8	26.5	25.2	2.5	1.6
10	1	3.5	3.0	21.7	23.9	35.3	26.5	0.5	4.3
	2	3.2	3.5	27.0	32.5	18.0	18.2	2.0	2.1
11	1	3.5	2.5	11.0	18.4	27.0	33.5	5.1	3.9
	2	2.0	2.8	16.4	8.1	35.4	29.3	2.1	5.0

表 14　例 2:水平 6 测试结果间极差

实验室	样本	测试结果间极差 %	k 统计量
1	1	2.6	0.624
	2	0.1	0.024
2	1	1.1	0.264
	2	2.5	0.600
3	1	7.6	1.825
	2	1.4	0.336
4	1	4.0	0.960
	2	8.1	1.945
5	1	1.3	0.312
	2	1.8	0.432
6	1	4.4	1.056
	2	2.1	0.504
7	1	3.9	0.936
	2	1.2	0.288
8	1	1.6	0.384
	2	1.1	0.264
9	1	0.6	0.144
	2	4.6	1.104
10	1	2.2	0.528
	2	5.5	1.320
11	1	7.4	1.777
	2	8.1	1.945

表 15　例 2:水平 6 样本间极差

实验室	样本间极差 %	k 统计量
1	6.75	1.767
2	4.40	1.152
3	1.00	0.262
4	2.25	0.589
5	2.05	0.537
6	2.55	0.668
7	3.15	0.825
8	3.35	0.877
9	1.70	0.445
10	6.95	1.819
11	2.55	0.668

表 16　例 2:水平 6 的单元平均值

实验室	单元平均值 %	h 统计量
1	26.425	1.475
2	13.750	−1.043
3	21.000	0.397
4	17.075	−0.382
5	13.425	−1.108
6	21.225	0.442
7	23.675	0.929
8	14.475	−0.899
9	18.250	−0.149
10	26.275	1.445
11	13.425	−1.108

表 17　例 2:根据表 13 中所有 8 个水平数据计算的平均值、极差平方和及标准差（剔除了有缺失数据的单元）

水平 j	实验室数 p'	总平均值 y_j %	极差平方和		标准差			
			SS_{rj} %^{2}	SS_{Hj} %^{2}	s_{xj} %	s_{rj} %	s_{Rj} %	s_{Hj} %
3	11	3.7	82.99	96.372 5	2.62	1.37	2.56	1.85
5	11	4.0	34.70	11.255 0	1.88	0.89	2.01	0.34
8	10	4.1	155.39	29.422 5	3.49	1.97	3.92	0.00
2	10	5.0	83.51	25.237 5	1.95	1.44	2.29	0.47
4	11	8.2	131.07	23.577 5	3.10	1.73	3.47	0.00
6	11	19.0	381.66	160.530 0	5.03	2.95	5.51	1.72
7	11	36.5	636.19	305.477 5	7.28	3.80	7.78	2.58
1	10	67.4	529.71	92.922 5	6.23	3.84	7.05	0.00

表 18　例 2:柯克伦和格拉布斯统计量

水平 j	实验室数 p'	测试结果间极差的柯克伦统计量	样本间极差的柯克伦统计量
3	11	0.203	0.664* (1)
5	11	0.461** (6)	0.374
8	10	0.298	0.465
2	10	0.232	0.238
4	11	0.169	0.550
8	11	0.172	0.301
7	11	0.157	0.536
1	10	0.237	0.680* (6)

单元平均值的格拉布斯统计量

水平 j	实验室数 p'	一个最小值	两个最小值	两个最大值	一个最大值
3	11	0.970	0.791	0.098** (1;6)	2.219
5	11	1.396	0.709	0.302	2.266
8	10	0.849	—	—	2.643** (6)
2	10	1.259	0.614	0.466	1.713
4	11	1.290	0.681	0.294	2.082
6	11	1.108	0.700	0.479	1.475
7	11	1.649	0.562	0.453	1.875
1	10	1.808	0.345	0.590	1.476

表 18(续)

注：括号内的数是指产生了歧离值或离群值的实验室号，其临界值如下：

统计检验	适用于	实验室数	相应于 GB/T 6379.2 中表的行数值	歧离值	离群值
		p'	p	(*)	(**)
柯克伦检验	测试结果间极差	10	20	0.389	0.480
		11	22	0.365	0.450
柯克伦检验	样本间极差	10	10	0.602	0.718
		11	11	0.570	0.684
对一个离群值的格拉布斯检验	单元平均值	10	10	2.290	2.482
		11	11	2.355	2.564
对两个离群值的格拉布斯检验	单元平均值	10	10	0.186 4	0.115 0
		11	11	0.221 3	0.144 8

5.9 非均匀物料设计计算的一般公式

在每个水平 j 上计算以下统计量：

a) 总平均值(对 i、t 和 k 求和)：

$$m_j = \sum\sum\sum y_{ijtk}/n_j \qquad (39)$$

其中 n_j 是求和中的测试结果数。

b) 对每个 i，实验室效应(对 t 和 k 求和)：

$$B_{ij} = \sum\sum(y_{ijtk} - m_j)/n_{ij}$$

$$= \text{实验室平均值} - \text{总平均值} \qquad (40)$$

其中 n_{ij} 是求和中的测试结果数。

c) 对 i 和 t，样本效应(对 k 求和)：

$$H_{ijt} = \sum(y_{ijtk} - m_j - B_{ij})/n_{ijt}$$

$$= \text{样本平均值} - \text{实验室平均值} \qquad (41)$$

其中 n_{ijt} 是求和中的测试结果数。

d) 对每一 i、t 和 k，残差：

$$z_{ijtk} = y_{ijtk} - m_j - B_{ij} - H_{ijt}$$

$$= \text{测试结果} - \text{样本平均值} \qquad (42)$$

e) 实验室平方和(对 i 求和)：

$$SS_{Lj} = \sum n_{ij} B_{ij}^2 \qquad (43)$$

f) 样本平方和(对 i 和 t 求和)：

$$SS_{Hj} = \sum\sum n_{ijt} B_{ijt}^2 \qquad (44)$$

g) 重复性平方和(对 i、t 和 k 求和)：

$$SS_{rj} = \sum\sum\sum z_{ijtk}^2 \qquad (45)$$

h) 自由度：

$$\nu_{Lj} = p'_j - 1 \qquad \nu_{Hj} = g_j - p'_j \qquad \nu_{rj} = n_j - g_j \qquad (46)$$

其中：

p'_j 是至少报告一个测试结果的实验室数；

g_j 是至少有一个测试结果被报告的样本数；

n_j 是测试结果总数。

i) 对每个 i，系数（对 t 求和）：

$$n_{ij} = \sum n_{ijt} \qquad \cdots\cdots (47)$$

$$K_{ij} = \sum n_{ijt}^2 \qquad \cdots\cdots (48)$$

j) 系数（对 i 求和）：

$$K_j = \sum n_{ij}^2 \qquad \cdots\cdots (49)$$

$$K'_j = \sum K_{ij} \qquad \cdots\cdots (50)$$

$$K''_j = \sum K_{ij}/n_{ij} \qquad \cdots\cdots (51)$$

k) 重复性标准差 s_{rj}，样本间标准差 s_{Hj}，实验室间标准差 s_{Lj} 和再现性标准差 s_{Rj}：

$$s_{rj}^2 = SS_{rj}/\nu_{rj} \qquad \cdots\cdots (52)$$

$$s_{Hj}^2 = [SS_{Hj} - \nu_{Hj} \times s_{rj}^2]/(n_j - K''_j) \qquad \cdots\cdots (53)$$

$$s_{Lj}^2 = [SS_{Lj} - (K''_j - K'_j/n_j) \times s_{Hj}^2 - \nu_{Lj} \times s_{rj}^2]/(n_j - K_j/n_j) \qquad \cdots\cdots (54)$$

$$s_{Rj}^2 = s_{rj}^2 + s_{Lj}^2 \qquad \cdots\cdots (55)$$

注：上述公式根据 Scheffé 建立的统计理论推导而得的。

5.10 例 3：一般公式的应用

5.10.1 例 2 中水平 4 的数据为一般公式的应用提供了一个例子，某些测试结果被省略了（见表 19）。用 5.9 的公式计算的总平均值列于表 19，平方和、自由度及各系数则分别列于表 20，表 21 和表 22 中。

5.10.2 应用 5.9 中第 k）步的式（52）至式（55），得：

$$s_{rj}^2 = SS_{rj}/\nu_{rj} = 36.895\,0/16$$

因此

$$s_{rj} = 1.52(\%)$$

同时得到

$$s_{Hj}^2 = [SS_{Hj} - \nu_{Hj} \times s_{rj}^2]/(n_j - K''_j) = [29.907\,5 - 9 \times 1.518\,5^2]/(36 - 19.667)$$

因此

$$s_{Hj} = 0.75(\%)$$

又

$$s_{Lj}^2 = [SS_{Lj} - (K''_j - K'_j/n_j) \times s_{Hj}^2 - \nu_{Lj} \times s_{rj}^2]/(n_j - K_j/n_j) = [378.853\,1 - (19.666\,7 - 68/36) \times 0.748\,7^2 - 10 \times 1.518\,5^2]/(36 - 130/36)$$

因此

$$s_{Lj} = 3.27(\%)$$

$$s_{Rj} = \sqrt{1.52^2 + 3.27^2} = 3.61(\%)$$

表 19 例 3：硫酸镁坚固性试验水平 4 的测试结果

实验室 i	样本 t	测试结果 $k=1$ %	测试结果 $k=2$ %
1	1	—	10.1
	2	13.9	13.8
2	1	—	—
	2	8.3	4.8

表 19(续)

实验室 i	样本 t	测试结果 $k=1$ %	测试结果 $k=2$ %
3	1	—	7.0
	2	—	12.0
4	1	9.4	—
	2	—	—
5	1	3.7	6.3
	2	3.3	3.7
6	1	16.5	12.3
	2	13.2	16.8
7	1	7.5	9.3
	2	11.1	8.3
8	1	5.7	6.8
	2	4.8	5.5
9	1	6.6	7.0
	2	4.9	6.3
10	1	5.5	5.8
	2	3.9	5.7
11	1	9.5	7.2
	2	8.1	7.4
总平均值：$m_j=8.111\ 1$ 测试结果数：$n_j=36$			

表 20 例 3：实验室平方和的计算

实验室	实验室平均值 %	测试结果数 n_{ij}	实验室效应 B_{ij} %	系数 K_{ij}
1	12.600	3	4.488 9	5
2	6.550	2	−1.561 1	4
3	9.500	2	1.388 9	2
4	9.400	1	1.288 9	1
5	4.250	4	−3.861 1	8
6	14.700	4	6.588 9	8
7	9.050	4	0.938 9	8
8	5.700	4	−2.411 1	8
9	6.200	4	−1.911 1	8
10	5.225	4	−2.886 1	8
11	8.050	4	−0.061 1	8
实验室平方和：$SS_{Lj}=378.853\ 1\%^2$ 实验室自由度：$\nu_{Lj}=11-1=10$ 系数：$K_j=130$ $K'_j=68$ $K''_j=19.666\ 7$				

表 21 例 3:样本平方和的计算

实验室 i	样本 t	样本平均值 %	测试结果数 n_{ijt}	样本效应 H_{ijt} %
1	1 2	10.10 13.85	1 2	−2.500 1.250
2	1 2	— 6.55	0 2	— 0.000
3	1 2	7.00 12.00	1 1	−2.500 2.500
4	1 2	9.40 —	1 0	0.000 —
5	1 2	5.00 3.50	2 2	0.750 −0.750
6	1 2	14.40 15.00	2 2	−0.300 0.300
7	1 2	8.40 9.70	2 2	−0.650 0.650
8	1 2	6.25 5.15	2 2	0.550 −0.550
9	1 2	6.80 5.60	2 2	0.600 −0.600
10	1 2	5.65 4.80	2 2	0.425 −0.425
11	1 2	8.35 7.75	2 2	0.300 −0.300
样本平方和:$SS_{H_{ij}}=29.907\ 5\%^2$ 样本自由度:$\nu_{H_{ij}}=20-11=9$				

表 22 例 3:重复性平方和的计算

实验室 i	样本 t	测试结果 $k=1$ %	测试结果 $k=2$ %
1	1 2	— 0.05	0.00 −0.05
2	1 2	— 1.75	— −1.75
3	1 2	— —	0.00 0.00
4	1 2	0.00 —	— —
5	1 2	−1.30 −0.20	1.30 0.20

表 22(续)

实验室 i	样本 t	测试结果 $k=1$ %	测试结果 $k=2$ %
6	1 2	2.10 −1.80	−2.10 1.80
7	1 2	−0.90 1.40	0.90 −1.40
8	1 2	−0.55 −0.35	0.55 0.35
9	1 2	−0.20 −0.70	0.20 0.70
10	1 2	−0.15 −0.90	0.15 0.90
11	1 2	1.15 0.35	−1.15 −0.35
重复性平方和：$SS_r=36.895$ 重复性自由度：$\nu_r=36-20=16$			

6 数据分析的稳健方法

6.1 数据分析稳健方法的应用

6.1.1 在 GB/T 6379.2—2004 中，建议对精密度试验得到的数据进行两种离群值的检验(柯克伦检验和格拉布斯检验)。如果两种检验中的一个或两个统计量超过 1% 显著水平的临界值时，则应剔除相应的数据(除非统计专家有充分的理由认为应将其保留)。在实际应用中，这样做通常并不容易。考虑 4.8 例 1 中的离群值检验的结果，这些结果在表 8 中给出。根据格拉布斯检验，实验室 5(在水平 10)仅仅有一个单元平均值由于非常极端应被判为离群值，但在其他水平还有 3 个歧离值，并且图 3 也明显表明该实验室异常。在此情形，统计专家必须要在下列三种可能决定中进行取舍：

a) 保留实验室 5 的所有数据；

b) 剔除实验室 5 水平 10 的数据；

c) 剔除实验室 5 的所有数据。

统计专家的决定将会对重复性和再现性标准差值的计算产生重大影响。常会遇到某些精密度试验数据处于歧离值和离群值之间边缘附近的数据，而对此做出的判断(剔除或保留)，对最后的计算结果有很大影响。这样往往不会令人满意。本章介绍的稳健方法允许对数据不需要作出取舍的判定从而不会影响计算结果的情况下，来对数据进行分析。因此若有理由认为精密度试验结果可能包含有离群值，则最好使用稳健方法。

6.1.2 GB/T 6379.2—2004 第 5 章中讨论的基本模型，假定应用同一种测量方法的所有实验室均可确定一个相同的重复性标准差的数值。实际上，通常的情形是某些实验室的重复性会比其他实验室差些，例如 5.8 中例 2 的图 5 所示。在这个试验中，实验室 6 显然比实验室 9 的重复性要差的多。因此在本例中，所有实验室具有相似重复性的假定显然不成立。当某些参与精密度试验的实验室第一次使用这种测量方法，或对此还缺少经验时，它的重复性就会较差，此时特别适宜采用稳健方法。

6.1.3 使用稳健方法[5]的目的，是在对精密度试验的数据进行分析时，所计算的重复性与再现性标准差的值不受离群数据的影响。如果将参与试验的实验室分为两类：一类数据质量高，另一类数据质量

低，则稳健方法所得的重复性标准差与再现性标准差的值应对高质量数据的那些实验室是有效的，而又不受那些低质量数据实验室的影响(只要低质量数据实验室不是很多)。

6.1.4 数据分析稳健方法的使用，不影响精密度试验的计划、组织与执行。应由统计专家决定是采用稳健方法还是采用剔除离群值的方法，然后向领导小组报告。在使用稳健方法时，仍应将 GB/T 6379.2 或 GB/T 6379.5 中描述的离群值检验和一致性检查用于这些数据，且对出现的任何离群值的原因及 h 和 k 统计量的表示模式进行检查。不过不管检验和检查的结果如何，数据都不予以剔除。

6.1.5 h 和 k 统计量的分母是用所报告的数据按 GB/T 6379.2—2004 描述的计算方法得出的标准差。若数据中含有离群值，则这些离群值会使分母变大，从而扭曲了相应的统计量的图形。例如，如果在试验的某个水平上，试验室的一个单元平均值是离群值，且比同水平的任何其他离群值大很多，则在那个水平的 h 统计量图中将会显示相应的 h 值异常的大，以至于即使其他实验室也有离群值，但相应的 h 统计量都将很小。对总平均值计算 h 统计量时也可会发生类似情况。若在 h 和 k 统计量中以标准差的稳健估计为分母，计算 h 统计量时也用总平均值的稳健估计，就不会发生这种扭曲。这就是推荐使用稳健方法的目的。

6.1.6 用精密度试验数据可计算以下两类统计量：

a) 单元平均值，由此计算用于度量实验室间变异的标准差；

b) 单元内标准差或极差(或在分割水平时的差值)，联合起来用于度量实验室内的变异。

这里介绍的稳健方法并不取代上述单元平均值、标准差、极差或差值，而是将这些量组合起来，提供另一种用以计算重复性与再现性标准差的统计量。

例如，对 GB/T 6379.2—2004 中的均匀水平设计中的某个水平的数据，分析的第一步是计算每个单元测试结果的平均值和标准差。然后用单元平均值计算用以度量实验室间变异的标准差。按本章的稳健方法，用“算法 A”计算该标准差，且在计算过程中那些应用格拉布斯检验被判为离群值的单元平均值并不被剔除。在这个设计中，联合各单元标准差来估计重复性标准差。而按稳健分析，用“算法 S”计算该估计值，且对那些应用柯克伦检验被判为离群值的单元标准差也不被剔除。两种方法的任何一种(GB/T 6379.2—2004 或这里所描述的方法)，都可以按相同的方式计算重复性和再现性标准差的估计值。

一个更复杂的例子是 GB/T 6379.3 附录 C 中给出的六因素错层套设计。根据这个设计，分析的第一步是计算每个实验室(在每个水平)数据的平均值 $y_{i(1)},\cdots,y_{i(5)}$ 以及一系列极差 $w_{i(1)},\cdots,w_{i(5)}$，这些极差包含由试验中所检查的各因素引起的变异的信息。应用稳健方法对数据进行分析，将“算法 A”用于单元平均值，将“算法 S”依次用于每组极差。利用这些计算所得到的统计量，按 GB/T 6379.3 描述的相同的分析方法即可获得重复性、中间精密度及再现性标准差的估计值。

6.1.7 之所以选择包含在 GB/T 6379 本部分的稳健方法，是因为它们可以用于 GB/T 6379 的第 2、3、4 和 5 部分给出的所有试验设计，也因为它们的计算相对简单。然而应注意到这些稳健方法提供的是一种从单元算术平均值和单元标准差出发，以稳健方式将它们相结合，而不是以稳健方式将单个测试结果结合。有一些稳健方法是以某种稳健方式将把单元内的测试结果相结合，这些方法在实际应用中较为复杂。

6.2 稳健分析：算法 A

6.2.1 对于所使用的数据，本算法可以求得其平均值与标准差的稳健值，它可应用于：

a) 任何设计的单元平均值；

b) 分割水平设计的单元差值。

6.2.2 记 $x_{(1)},x_{(2)},\cdots,x_{(i)},\cdots,x_{(p)}$ 为按递增顺序排列的 p 个数据，以 x^* 和 s^* 表示这些数据的稳健平均值和稳健标准差。

6.2.3 计算 x^* 和 s^* 的初始值如下(med 表示中位数)：

$$x^* = \operatorname{med} x_{(i)} \qquad (i = 1,2,\cdots,p) \qquad \cdots\cdots(56)$$

$$s^{*} = 1.483 \times \text{med} \mid x_{(i)} - x^{*} \mid (i = 1,2,\cdots,p) \qquad (57)$$

6.2.4 更新 x^{*} 和 s^{*} 的值如下：令

$$\varphi = 1.5s^{*} \qquad (58)$$

对每个 $x_{(i)}(i=1,2,\cdots,p)$，计算

$$x_{(i)}^{*} = \begin{cases} x^{*} - \varphi, x_{(i)} < x^{*} - \varphi \\ x^{*} + \varphi, x_{(i)} > x^{*} + \varphi \\ x_{(i)}, \text{其他} \end{cases} \qquad (59)$$

计算新的 x^{*} 和 s^{*} 值：

$$x^{*} = \sum_{i=1}^{p} x_{(i)}^{*}/p \qquad (60)$$

$$s^{*} = 1.134\sqrt{\sum_{i=1}^{p}(x_{(i)}^{*} - x^{*})^2/(p-1)} \qquad (61)$$

6.2.5 通过迭代计算，即重复6.2.4中的计算若干次，直到所得 x^{*} 和 s^{*} 新的估计值的变化很小为止。在计算机编程计算中，这极为简单。

6.2.6 一种可手算不需要迭代，从而更易于应用的方法，是将6.2.4中的式(60)和式(61)改写为：

$$x^{*} = x' + 1.5 \times (u_{U} - u_{L})s^{*}/(p - u_{L} - u_{U}) \qquad (62)$$

$$(s^{*})^2 = (p - u_{L} - u_{U} - 1) \times (s')^2/[(p-1)/1.134^2 - 1.5^2(pu_{L} + pu_{U} - 4u_{L}u_{U})/(p - u_{L} - u_{U})] \qquad (63)$$

其中：

u_{L} 是满足 $x_{(i)} < x^{*} - \phi$ 的数据 $x_{(i)}$ 个数；

u_{U} 是满足 $x_{(i)} > x^{*} + \phi$ 的数据个数；

x' 和 s' 是满足 $|x_{(i)} - x^{*}| \leqslant \phi$ 的 $(p - u_{L} - u_{U})$ 个数据 $x_{(i)}$ 的平均值和标准差。

若 u_{L} 和 u_{U} 已知，即可直接用来计算 x^{*} 和 s^{*}。一种方法是按一定顺序试算各种可能(如按 $u_{L}=0$，$u_{U}=0$；$u_{L}=0$，$u_{U}=1$；$u_{L}=1$，$u_{U}=0$；$u_{L}=1$，$u_{U}=1$ 等顺序试算)，直到求得一个有效解。所谓有效解是距 x^{*} 超过 $1.5s^{*}$ 的实际数据个数正好等于用于计算 x^{*} 和 s^{*} 的 u_{L} 和 u_{U}。实际上，分析者可利用图4那样的条形图帮助识别那些距 x^{*} 可能超过 $1.5s^{*}$ 的数据，以便通过少量的试算求得有效解。

另一个可能方法是用迭代法来求出一个近似解，然后解式(62)和式(63)得到精确解。下面的例子所用的就是这种方法。

6.3 稳健分析：算法S

6.3.1 本算法用于任何设计的实验室内标准差(或实验室内极差)，用该算法可得到标准差或极差的一个稳健的联合值。

6.3.2 记 $w_{(1)}, w_{(2)}, \cdots, w_{(i)}, \cdots, w_{(p)}$ 为按递增顺序排列的 p 个数据(极差或标准差数据)，以 w^{*} 表示稳健联合值，ν 表示与每个 $w_{(i)}$ 相关的自由度(当 $w_{(i)}$ 是极差时，$\nu=1$；当 $w_{(i)}$ 是 n 个结果的标准差时，$\nu=n-1$)。

从表23查得本算法所需的 ξ 和 η 的值。

6.3.3 计算 w^{*} 的初始值如下(med 表示中位数)：

$$w^{*} = \text{med}\, w_{(i)} (i = 1,2,\cdots,p) \qquad (64)$$

6.3.4 更新 w^{*} 的值如下：令

$$\Psi = \eta \times w^{*} \qquad (65)$$

对每个 $w_{(i)}(i=1,2,\cdots,p)$，计算

$$w_{(i)}^{*} = \begin{cases} \Psi, w_{(i)} > \Psi \\ w_{(i)}, \text{其他} \end{cases} \qquad (66)$$

计算新的 w^* 值：

$$w^* = \xi\sqrt{\sum_{i=1}^{p}(w_{(i)}^*{}^2/p} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(67)$$

6.3.5 通过6.3.4的迭代计算即可得到稳健估计值 w^*，即重复6.2.4中的计算数次，直到 w^* 两次估计值的变化很小为止。这在计算机上通过简单的编程即可实现。

6.3.6 与6.2.6类似，一种不需要迭代可手算的，从而也是更易应用的方法，是将6.3.4中的(67)式改写为：

$$(w^*)^2 = [\xi^2/p]\times[\sum{}'(w_{(i)}^*)^2 + u_U\times(\eta w^*)^2] \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots(68)$$

其中：

$\sum{}'$表示对满足 $w_{(i)}\leqslant\Psi$ 的 $w_{(i)}$ 求和；

u_U 是满足 $w_{(i)}>\Psi$ 的 $w_{(i)}$ 的个数。

这可以通过依次对 $u_U=0$，$u_U=1$，$u_U=2$，等等进行试算，直到得到一个有效解。有效解是超过 $\eta\times w^*$ 的 $w_{(i)}$ 的实际个数恰为 u_U。实际上分析者可利用图4那样的条形图帮助识别那些可能超过 $\eta\times w^*$ 的极差，以便通过少量的试算求得有效解。

下面例子中所用的方法是先用迭代法求出一个近似解，然后解式(68)求得精确解。

表23 算法S：稳健分析所需的系数

自由度 ν	限系数 η	修正系数 ξ
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
注——ξ 和 η 数值的推导见附录B。		

6.4 公式：均匀水平设计特定水平的稳健分析

6.4.1 在均匀水平设计中，对某特定水平，将算法S应用于单元极差或单元标准差，计算该水平的重复性标准差的估计值 s_r，进而根据6.3.4中的(67)式求得稳健值 w^*。如将算法S用于单元标准差，则

$$s_r = w^* \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(69)$$

如果每个单元有两个测试结果，且将算法S用于单元极差，则

$$s_r = w^*/\sqrt{2} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(70)$$

6.4.2 某一水平单元平均值的标准差的稳健估计值 s_d 可按以下方法得到：将算法A应用于单元平均值，根据6.2.4中的式(61)求得稳健值 s^*，

$$s_d = s^* \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(71)$$

6.4.3 接下来可用下式确定试验室间标准差 s_L：

$$s_L = \sqrt{(s_d^2 - s_r^2)/n} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(72)$$

其中 n 是每个单元的测试结果数。

若根号内的表达式为负值，则置

$$s_L = 0 \quad \cdots\cdots(73)$$

该水平的再现性标准差计算如下：

$$s_R = \sqrt{s_L^2 + s_r^2} \quad \cdots\cdots(74)$$

6.5 例 4：均匀水平设计特定水平的稳健分析

6.5.1 GB/T 6379.2—2004 的例 3 是一个数据包含歧离值和离群值的例子。该例的水平 5 令人关注，因为根据格拉布斯检验，试验室 1 的一个单元平均值很接近歧离值；而根据科克伦检验，实验室 6 的一个单元极差也很接近歧离值。表 24 重列了有关数据。

6.5.2 若保留所有试验室的数据，用 GB/T 6379.2—2004 中 7.4 的有关公式，可计算重复性和再现性标准差的估计：

$$p = 9$$
$$m = 20.511$$
$$s_r = 0.585$$
$$s_d = 1.727$$
$$s_L = 1.677$$
$$s_R = 1.776$$

6.5.3 然而，根据 GB/T 6379.2—2004，数据分析员基于从试验中其他水平获得的信息以及对实验室 6 受试样本的怀疑，认为在计算中应剔除实验室 1 与实验室 6 的数据，由此得到：

$$p = 7$$
$$m = 20.412$$
$$s_r = 0.393$$
$$s_d = 0.573$$
$$s_L = 0.51$$
$$s_R = 0.637$$

显然，剔除两个实验室数据的决定对于重复性和再现性标准差的估计有本质的影响。

6.5.4 分析的第一步是获得重复性标准差的一个稳健估计。计算可以方便的用表 25 表示，其中单元极差按递增顺序排列。用算法 S 的迭代法，所得的结果列于表 25。本例中的单元极差的自由度 $\nu=1$，所以 $\xi=1.097$，$\eta=1.645$。从表中所示的 4 次迭代，得 $w^*=0.7$，且只有一个单元极差（$w_9^*=1.98$）超过 Ψ。若在计算机上进行计算，迭代过程应继续到连续得到的两个 w^* 的值相差很小为止。

也可以直接按以下过程求解，在 6.3.6 中的式(68)中，将

$$u_U = 1$$

$$\sum_{i=1}^{p}(w_{(i)}^*)^2/p = 0.249\ 5$$

代入 6.3.6 中的(68)式，得到

$$(w^*)^2 = 1.097^2 \times 0.249\ 5 + (1.097 \times 1.645 w^*)^2/9$$

（若假定 $u_U=1$ 是正确的），则解为：

$$w^* = 0.69(\%)$$

由 $w^*=0.69$ 可以给出 $\Psi=1.645\times0.69=1.14$，所以只有 w_9^* 超过了 Ψ；下一步用 1.14 代替 $w_{(9)}^*$，再次得到 $w^*=0.63\times1.907=0.69$，从而验证了解的有效性。

因此重复性标准差的估计值为：

$$s_r = 0.69/\sqrt{2} = 0.49(\%)$$

其值介于 6.5.2 和 6.5.3 中给出的两个估计值之间。

6.5.5 分析的下一步是确定单元平均值标准差的稳健估计值。将算法 A 应用于单元平均值，得到表 26 的结果，其中单元平均值按递增顺序排列。从表中所示的 4 次迭代，稳健值是 $x^*=20.412$，$s^*=1.1$。只有两个极端单元平均值（$x^*_{(1)}=17.570$ 和 $x^*_{(9)}=24.140$）距 x^* 超过 ϕ。若在计算机上进行计算，迭代过程应继续到连续得到的两组 x^* 和 s^* 的值相差很小为止。

如果进行手算，数据分析者应按 6.2.6 中描述的直接方法，先试用：

$$u_L = u_U = 1$$

得到

$$x' = 20.412\%$$
$$s' = 0.573\%$$

因此，从 6.2.6 的式(62)和式(63)有：

$$(s^*)^2 = 6\times(0.573)^2/(8/1.134\,2^2 - 1.5^2(9+9-4)/7)$$

由此

$$s^* = 1.070(\%)$$
$$x^* = x' = 20.412(\%)$$

由 s^* 的值，得出 $\phi=1.605$（故如同所假定的，只有 $x^*_{(1)}$ 和 $x^*_{(9)}$ 距 x^* 超过了 ϕ）。用 18.807 代替 $x^*_{(1)}$，22.017 代替 $x^*_{(9)}$，再次得到 $s^*=0.944\times1.134=1.070$，从而验证了解的有效性。

因此根据 6.4.3 的式(72)，实验室间标准差的估计值是：

$$s_L = \sqrt{1.070^2 - (0.49^2/2)} = 1.012(\%)$$

而根据 6.4.3 的式(74)，再现性标准差的估计值是：

$$s_R = \sqrt{1.012^2 + 0.49^2} = 1.124(\%)$$

此值也介于 6.5.2 和 6.5.3 中给出的估计值之间。

表 24 例 4：木馏油含量的热滴定（%）

实验室 i	数据 %		单元平均值 %	单元极差 %
1	24.28	24.00	24.140	0.28
2	20.40	19.91	20.155	0.49
3	19.30	19.70	19.500	0.40
4	20.30	20.30	20.300	0.00
5	20.53	20.88	20.705	0.35
6	18.56	16.58	17.570	1.98
7	19.70	20.50	20.100	0.80
8	21.10	20.78	20.940	0.32
9	20.71	21.66	21.185	0.95

表 25 例 4：用算法 S 计算单元极差（木馏油，%）
（$\nu=1$；$\xi=1.097$；$\eta=1.645$）

迭代	0[1]	1	2	3	4
Ψ	—	0.66	0.86	1.00	1.09
w_1^*	0.00	0.00	0.00	0.00	0.00
w_2^*	0.28	0.28	0.28	0.28	0.28

表 25(续)

迭代	0[1)]	1	2	3	4
w_3^*	0.32	0.32	0.32	0.32	0.32
w_4^*	0.35	0.35	0.35	0.35	0.35
w_5^*	0.40	0.40	0.40	0.40	0.40
w_6^*	0.49	0.49	0.49	0.49	0.49
w_7^*	0.80	0.66	0.80	0.80	0.80
w_8^*	0.95	0.66	0.86	0.95	0.95
w_9^*	1.98	0.66	0.86	1.00	1.09
联合 w 新 w^*	0.83 0.40[2)]	0.47 0.52	0.56 0.61	0.60 0.66	0.62 0.68

1) 迭代 0 系列将表 24 数据按递增顺序重新排列。

2) 0.40 是极差的中位数(见 6.3.3 式(64))。

表 26 例 4:用算法 A 计算单元平均值(木馏油,%)

迭代	0[1)]	1	2	3	4
φ	—	1.424	1.478	1.514	1.539
$x^*-\varphi$ $x^*+\varphi$	— —	18.876 21.724	18.909 21.865	18.893 21.921	18.872 21.950
x_1^*	17.570	18.876	18.909	18.893	18.872
x_2^*	19.500	19.500	19.500	19.500	19.500
x_3^*	20.100	20.100	20.100	20.100	20.100
x_4^*	20.155	20.155	20.155	20.155	20.155
x_5^*	20.300	20.300	20.300	20.300	20.300
x_6^*	20.705	20.705	20.705	20.705	20.705
x_7^*	20.940	20.940	20.940	20.940	20.940
x_8^*	21.185	21.185	21.185	21.185	21.185
x_9^*	24.140	21.724	21.865	21.921	21.950
平均值	20.511	20.387	20.407	20.411	20.412
标准差	1.727	0.869	0.890	0.905	0.916
新 x^* 新 s^*	20.300[2)] 0.949[2)]	20.387 0.985	20.407 1.009	20.411 1.026	20.412 1.039

1) 迭代 0 系列将表 24 数据按递增顺序重新排列。

2) 是根据 6.2.3 式(56)和式(57)得到。

6.6 公式:分割水平设计特定水平的稳健分析

6.6.1 分割水平设计的某一水平重复性标准差 s_r 的稳健估计值可以通过将算法 A 用于该水平的单元差值,按式(61)求得稳健值 s^*,然后利用下式计算 s_r:

$$s_r = s^* / \sqrt{2} \qquad \cdots\cdots (75)$$

6.6.2 对某一特定水平,单元平均值标准差 s_y 的稳健估计值也可通过将算法 A 用于该水平的单元平均值,按式(61)求得稳健值 s^*,然后利用下式计算 s_y:

$$s_y = s^* \qquad \cdots\cdots(76)$$

4.5.6 中给出的公式可用于计算该水平再现性标准差的估计值。

6.7 例 5:分割水平设计特定水平的稳健分析

6.7.1 4.8 中例 1 的数据包含若干歧离值和一个离群值(见表 8)。图 3 也表明实验室 5 的结果有一致的负偏倚。如果数据分析者不能找到造成这些异常的原因,他就会很为难;如果有数据需要剔除的话,必须决定在计算重复性和再现性标准差时应该剔除哪些数据。下面用水平 14 的数据(参见表 4)说明根据稳健分析所得到的结果。

6.7.2 为得到重复性标准差的稳健估计值,将算法 A 应用于单元差值(表 5),结果列于表 27 中,其中单元差值按递增顺序排列。经表中所示的 4 次迭代,稳健值 $x^* = 8.29$,$s^* = 0.36$,只有 $x^*_{(9)}$ 距 x^* 超过 φ。

令 $u_L = 0$,$u_U = 1$,应用 6.2.6 中所描述的方法,得:

$$x' = 8.219, s' = 0.257$$

故 6.2.6 中式(62)和式(63)可写为:

$$x^* = 8.219 + 1.5 \times s^*/8$$

$$(s^*)^2 = 7 \times (0.257)^2/[8/1.134^2 - 1.5^2(0 + 9 - 0)/8]$$

由此得到

$$s^* = 0.354$$

按 6.6.1 中的式(75),有:

$$s_r = 0.354/\sqrt{2} = 0.250$$

于是单元差值稳健平均值估计为:

$$x^* = 8.219 + 1.5 \times 0.354/8 = 8.285$$

用这些 x^* 和 s^* 值,

$$\varphi = 1.5 \times 0.354 = 0.531$$

故

$$x^* - \varphi = 7.754(\%), x^* + \varphi = 8.816$$

在计算 x^* 和 s^* 时,我们假定了只有 x^*_9 超出上述界限。可证实情况的确如此,因此得到的是有效解。

6.7.3 将算法 A 用于单元平均值(根据表 6)给出表 28 所示的结果,其中单元平均值按递增顺序排列。此处的情况类似于表 26 遇到的情况,其中 $x^*_{(1)}$ 和 $x^*_{(9)}$ 距 x^* 超过 φ,x^* 收敛于 $x^*_{(2)},\cdots,x^*_{(8)}$ 的平均值 85.486。仍使用 6.2.6 中的方法,取 $u_L = u_U = 1$,则 $x^*_{(2)},\cdots,x^*_{(8)}$ 的平均值和标准差为:

$$x' = 85.486, s' = 0.209$$

因此根据 6.2.6 中的式(63),

$$(s^*)^2 = 6 \times (0.209)^2/[8/1.134^2 - 1.5^2(9 + 9 - 4)/7]$$

因而得:

$$s^* = 0.390$$

按 6.2.6 中的式(62),得:

$$x^* = 85.486$$

为检查解的有效性,计算得

$$\varphi = 1.5 \times 0.390 = 0.585$$

故

$$x^* - \varphi = 84.901(\%), x^* + \varphi = 86.071$$

如同假定的那样，将看到只有 $x_{(1)}^*$ 和 $x_{(9)}^*$ 落在了上述界限之外。

为求再现性标准差，用 6.6.2 中的式(76)，得到：

$$s_y = 0.309$$

然后由 4.5.6 中的式(13)，可得：

$$s_R = 0.410$$

因此，本例中稳健方法给出了估计值 s_r 和 s_R，比使用全部报告数据计算出的值(见表 7)要稍小一些。

表 27　例 5：用算法 A 计算单元差值(蛋白质，%)

迭代	0	1	2	3	4
φ	—	0.53	0.56	0.55	0.54
$x^*-\varphi$ $x^*+\varphi$	— —	7.85 8.91	7.74 8.86	7.74 8.84	7.75 8.83
x_1^*	7.81	7.85	7.81	7.81	7.81
x_2^*	7.93	7.93	7.93	7.93	7.93
x_3^*	8.13	8.13	8.13	8.13	8.13
x_4^*	8.14	8.14	8.14	8.14	8.14
x_5^*	8.38	8.38	8.38	8.38	8.38
x_6^*	8.40	8.40	8.40	8.40	8.40
x_7^*	8.44	8.44	8.44	8.44	8.44
x_8^*	8.52	8.52	8.52	8.52	8.52
x_9^*	9.31	8.91	8.86	8.84	8.83
平均值	8.340	8.300	8.290	8.288	8.287
标准差	0.436	0.326	0.322	0.317	0.315
新 x^* 新 s^*	8.380[1)] 0.356[1)]	8.300 0.370	8.290 0.365	8.288 0.359	8.287 0.357

1)　根据 6.2.3 式(56)和式(57)得到。

表 28　例 5：用算法 A 计算单元平均值(蛋白质，%)

迭代	0	1	2	3	4
φ	—	0.446	0.492	0.519	0.537
$x^*-\varphi$ $x^*+\varphi$	— —	85.104 85.996	85.009 85.993	84.971 86.009	84.950 86.024
x_1^*	84.525	85.104	85.009	84.971	84.950
x_2^*	85.140	85.140	85.140	85.140	85.140
x_3^*	85.345	85.345	85.345	85.345	85.345
x_4^*	85.385	85.385	85.385	85.385	85.385
x_5^*	85.550	85.550	85.550	85.550	85.550
x_6^*	85.575	85.575	85.575	85.575	85.575
x_7^*	85.660	85.660	85.660	85.660	85.660

表 28(续)

迭代	0	1	2	3	4
x_8^*	85.750	85.750	85.750	85.750	85.750
x_9^*	86.170	85.996	85.993	86.009	86.024
平均值	85.456	85.501	85.490	85.487	85.487
标准差	0.453	0.289	0.305	0.316	0.324
新 x^* 新 s^*	85.550[1)] 0.297[1)]	85.501 0.328	85.490 0.346	85.487 0.358	85.487 0.367

1) 根据 6.2.3 式(56)和式(57)得到。

6.8 公式:非均匀物料试验特定水平的稳健分析

6.8.1 对非均匀物料设计,在通常情况下 p'个实验室中的每个实验室在每一水平下都制备两个样本,对每个样本进行两次测试。为获得重复性和再现性标准差的稳健估计,需按以下程序 3 次使用算法 A 和 S:

a) 根据 6.3.4 将算法 S 应用于测试结果间极差,按式(67)得稳健值 w^*。再令

$$SS_r = 2p'(w^*)^2 \quad \cdots\cdots(77)$$

b) 将算法 S 应用于样本间极差,按式(67)得另一稳健值 w^*。令

$$SS_H = p'(w^*)^2 \quad \cdots\cdots(78)$$

c) 根据 6.2.4 将算法 A 应用于单元平均值,按式(61),得稳健值 s^*。令

$$s_y = s^* \quad \cdots\cdots(79)$$

上述计算可以如下面的例子所示,方便的以表格形式表示,其中第一列的极差或平均值的以递增顺序排列。

6.8.2 5.5 中给出的公式可用于计算重复性和再现性标准差以及用以度量样本间变异的标准差 s_H 的估计值。

6.9 例 6:非均匀物料试验特定水平的稳健分析

6.9.1 5.8 中例 2 的水平 6 的数据不包括任何离群值或歧离值,所以此处用它们来说明用稳健方式所获得的结果。

6.9.2 将算法 S 应用于测试结果间极差(表 14),所得到的结果列于表 29。自由度 $\nu=1$,故 $\eta=1.645$,$\xi=1.097$。数据项数 $p=2p'=22$。根据表中所示的 4 次迭代,得稳健值 $w^*=4.5$,$w^*_{(15)}$ 至 $w^*_{(22)}$ 诸值都超过 Ψ。具体计算过程如下($\sum'$ 和 u_U 如 6.3 中所定义):

$$u_U = 4$$

$$\sum{}'(w^*)^2/p = 137.92/22 = 6.269\ 1$$

因而 6.3.6 中的(68)式成为:

$$(w^*)^2 = 1.097^2 \times 6.269\ 1 + 4(1.097 \times 1.645w^*)^2/22$$

从而得:

$$w^* = 4.30(\%)$$

根据 w^* 的值,$\Psi=7.1$,$w^*_{(19)}$ 至 $w^*_{(22)}$ 的 4 个值超过了 Ψ,所以验证实了 w^* 是一个有效解。

从 6.8.1 中式(77),得:

$$SS_r = 22 \times 4.30^2 = 406.78$$

6.9.3 将算法 S 第二次应用于样本间极差(表 15),结果列于表 30。根据表中所示的 4 次迭代,稳健值 $w^*=4.0$,$w^*_{(10)}$ 和 $w^*_{(11)}$ 超过了 Ψ。利用 6.3 中所定义的 $\sum'$ 和 u_U,在下面情况下:

$$u_U = 2$$

$$\sum'(w_i^*)^2/p' = 66.665/11 = 6.0605$$

因而 6.3.6 中的式(68)成为：

$$(w^*)^2 = 1.097^2 \times 6.0605 + 2(1.097 \times 1.645w^*)^2/11$$

从而得：

$$w^* = 4.23$$

不幸的是，由于与此对应的 $\Psi=1.645\times4.23=6.96$，但 $w_{(10)}^*$ 和 $w_{(11)}^*$ 没有超过该值，所以它不是一个有效解。这表明可能要按 $u_U=1$ 或 $u_U=0$ 求解。

先试 $u_U=1$，得

$$\sum'(w_i^*)^2/p' = 112.2275/11 = 10.2025$$

此时式(68)为：

$$(w^*)^2 = 1.097^2 \times 10.2025 + (1.097 \times 1.645w^*)^2/11$$

$$w^* = 4.18$$

对应的 $\Psi=1.645\times4.18=6.88$，因而这是一个有效解，因为只有 $w_{(11)}^*$ 超过该值。

用 6.8.1 中的式(78)可得：

$$SS_H = 11 \times 4.18^2 = 192.20$$

6.9.4 将算法 A 应用于单元平均值(表 16)，得到的结果列于表 31。经过 2 次迭代，计算结果即收敛：$s^*=5.70$(没有任何 $x_{(i)}^*$ 距 x^* 大于 ϕ)。

用 6.8.1 中的式(79)，得到

$$s_y = 5.70$$

6.9.5 综合 6.9.2、6.9.3 和 6.9.4 中得到的结果及 5.5.5 中的式(29)至式(33)，有：

$$s_r^2 = 406.78/44$$

$$s_R^2 = 5.70^2 + (406.78 - 192.20)/44$$

$$s_H^2 = 192.20/22 - 406.78/88$$

所以

$$s_r = 3.04$$

$$s_R = 6.11$$

$$s_H = 2.03$$

在本例中，稳健方法给出的估计值 s_r、s_R 和 s_H 比使用全部报告数据计算的值(5.8.3 表 17 中所给)稍大。

表 29 例 6：用算法 S 计算测试结果间极差(%)

($\nu=1$；$\xi=1.097$；$\eta=1.645$)

迭代	0	1	2	3	4
Ψ	—	3.9	5.1	5.9	6.4
w_1^*	0.1	0.1	0.1	0.1	0.1
w_2^*	0.6	0.6	0.6	0.6	0.6
w_3^*	1.1	1.1	1.1	1.1	1.1
w_4^*	1.1	1.1	1.1	1.1	1.1
w_5^*	1.2	1.2	1.2	1.2	1.2
w_6^*	1.3	1.3	1.3	1.3	1.3
w_7^*	1.4	1.4	1.4	1.4	1.4

表 29(续)

迭代	0	1	2	3	4
w_8^*	1.6	1.6	1.6	1.6	1.6
w_9^*	1.8	1.8	1.8	1.8	1.8
w_{10}^*	2.1	2.1	2.1	2.1	2.1
w_{11}^*	2.2	2.2	2.2	2.2	2.2
w_{12}^*	2.5	2.5	2.5	2.5	2.5
w_{13}^*	2.6	2.6	2.6	2.6	2.6
w_{14}^*	3.9	3.9	3.9	3.9	3.9
w_{15}^*	4.0	3.9	4.0	4.0	4.0
w_{16}^*	4.4	3.9	4.4	4.4	4.4
w_{17}^*	4.6	3.9	4.6	4.6	4.6
w_{18}^*	5.5	3.9	5.1	5.5	5.5
w_{19}^*	7.4	3.9	5.1	5.9	6.4
w_{20}^*	7.6	3.9	5.1	5.9	6.4
w_{21}^*	8.1	3.9	5.1	5.9	6.4
w_{22}^*	8.1	3.9	5.1	5.9	6.4
新 w 新 w^*	4.17 2.35[1)]	2.80 3.07	3.29 3.61	3.55 3.89	3.70 4.06

1) 根据6.3.3中式(64)得到。

表 30 例 6:用算法 S 计算样本间极差(%)
($\nu=1$;$\xi=1.097$;$\eta=1.645$)

迭代	0	1	2	3	4
Ψ	—	4.19	5.43	6.10	6.45
w_1^*	1.00	1.00	1.00	1.00	1.00
w_2^*	1.70	1.70	1.70	1.70	1.70
w_3^*	2.05	2.05	2.05	2.05	2.05
w_4^*	2.25	2.25	2.25	2.25	2.25
w_5^*	2.55	2.55	2.55	2.55	2.55
w_6^*	2.55	2.55	2.55	2.55	2.55
w_7^*	3.15	3.15	3.15	3.15	3.15
w_8^*	3.35	3.35	3.35	3.35	3.35
w_9^*	4.40	4.19	4.40	4.40	4.40
w_{10}^*	6.75	4.19	5.43	6.10	6.45
w_{11}^*	6.95	4.19	5.43	6.10	6.45
新 w 新 w^*	3.82 2.55[1)]	3.01 3.30	3.38 3.71	3.58 3.92	3.69 4.05

1) 根据6.3.3中式(64)得到。

表 31 例 6:用算法 A 计算单元平均值(%)

迭代	0	1	2	3	4
φ		10.005	8.550		
$x^*-\varphi$ $x^*+\varphi$	— —	8.245 28.255	10.450 27.550		
x_1^*	13.425	13.425	13.425		
x_2^*	13.425	13.425	13.425		
x_3^*	13.750	13.750	13.750		
x_4^*	14.475	14.475	14.475		
x_5^*	17.075	17.075	17.075		
x_6^*	18.250	18.250	18.250		
x_7^*	21.000	21.000	21.000		
x_8^*	21.225	21.225	21.225		
x_9^*	23.675	23.675	23.675		
x_{10}^*	26.275	26.275	26.275		
x_{11}^*	26.425	26.425	26.425		
平均值	19.00	19.00	19.00		
标准差	5.03	5.03	5.03		
新 x^* 新 s^*	18.25[1)] 6.67[1)]	19.00 5.70	19.00 5.70		

1) 根据 6.2.3 中式(56)和式(57)得到。

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

$B_{(1)}, B_{(2)}, \cdots$ 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

$C_{\text{crit}}, C'_{\text{crit}}, C''_{\text{crit}}$ 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（标准物质/标准材料）

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(σ_R/σ_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

GB/T 6379.5 中增加的符号和缩略语

D 分割水平试验单元内差值

g 实验室在一个水平上的测试样本数

H 测量结果的样本随机误差分量

K 单元中测量结果数的函数

p' 参与实验间试验的实验室个数

SS 平方和

u_L 稳健分析中小于下限的数据个数

u_U 稳健分析中大于上限的数据个数

z 残差

Φ 标准差的比

φ 稳健分析中使用的限(算法 A)

η 稳健分析中的限系数(算法 S)

Ψ 稳健分析中使用的限(算法 S)

ξ 稳健分析中使用的修正系数(算法 S)

GB/T 6379.5 中增加的用于下标的符号

a,b 分割水平试验样本的标识

t 实验室 i 在水平 j 上样本的标识

H 样本间

GB/T 6379.5 中增加的用于上标的符号

* 稳健估计值

附 录 B
（资料性附录）
算法 A 和算法 S 中所用系数的推导

B.1 引言

对精密度试验数据进行稳健分析的方法是由英国皇家化学会分析方法委员会提出的（参见参考文献[6]）。本部分的算法 A 引自所列的论文。算法 A 中计算 s^* 的系数 1.134 也取自于所列的论文（用论文的记号，它是 $c=1.5$ 时 $1/\sqrt{\beta}$ 的值）。

算法 S 类似于参考文献[6]中给出的方法的一种特殊情形：每个实验室对每个水平都报告 $n=2$ 个测量。它提供了一种将稳健分析用于两个水平以上因素精密度试验（如 GB/T 6379 本部分第 5 章中非均匀物料设计及 GB/T 6379.3 中的错层套设计）的方便方法。算法 S 中所用的系数推导如下。

B.2 本附录中使用的符号

σ　标准差的真值

s　标准差 σ 的估计值

ν　s 的自由度

ω　$\nu+2$

ξ　算法 S 的修正系数

η　算法 S 的限系数

χ_ν^2　自由度为 ν 的 χ^2 变量

$$s^* = \begin{cases} s & 若\ s \leqslant \eta\sigma \\ \phi\sigma & 若\ s > \eta\sigma \end{cases}$$

B.3 限系数和修正系数的推导

修正系数 ξ 定义是，为使 $(s^*)^2$ 成为 σ^2 的无偏估计，需对 s^* 进行修正的值，即：

$$E\{(\xi \times s^*)^2\} = \sigma^2 \quad \cdots\cdots\cdots\cdots (\text{B.1})$$

上式可改写为：

$$E\{\nu(s^*/\sigma)^2\} = \nu/\xi^2 \quad \cdots\cdots\cdots\cdots (\text{B.2})$$

其中{ }中的量与 $\nu(s/\sigma)^2$ 密切相关，它是自由度为 ν 的 χ^2 变量。

χ^2 变量 χ_ν^2 的概率密度函数为：

$$f(x) = e^{-x/2} x^{(\nu/2-1)} 2^{-\nu/2} / \Gamma(\nu/2) \quad \cdots\cdots\cdots\cdots (\text{B.3})$$

故

$$E\{\nu(s^*/\sigma)^2\} = \int_0^{\nu\eta^2} x f(x)\mathrm{d}x + \int_{\nu\eta^2}^{\infty} \nu\eta^2 f(x)\mathrm{d}x \quad \cdots\cdots\cdots\cdots (\text{B.4})$$

因为限 $s \leqslant \eta\sigma$ 等价于 $\nu(s/\sigma)^2 \leqslant \nu\eta^2$，所以(B.4)式右边的第二项为：

$$\nu\eta^2 \times P(\chi_\nu^2 > \nu\eta^2) = \nu\eta^2 \times P(s > \eta\sigma) \quad \cdots\cdots\cdots\cdots (\text{B.5})$$

对于算法 S，限系数 η 的确定是使 $\eta\sigma$ 是 s 分布的上 10% 分位点，即

$$P(s > \eta\sigma) = 0.1 \quad \cdots\cdots\cdots\cdots (\text{B.6})$$

GB/T 4086.2—1983《统计分布数值表 χ^2 分布数值表》给出了 GB/T 6379 本部分表 23 所需的 η 值。从式(B.5)和式(B.6)知，式(B.4)右边的第二项为 $0.1\nu\eta^2$。注意 η 依赖于 s 的自由度。

式(B.4)右边的第一项可以写成：

$$\int_0^{\nu\eta^2} xe^{-x/2}x^{\nu/2-1}2^{-\nu/2}/\Gamma(\nu/2)\mathrm{d}x$$

根据 Γ 函数的一个熟知的性质($\omega=\nu+2$)：

$$\Gamma(\omega/2) = \Gamma(\nu/2+1) = (\nu/2)\times\Gamma(\nu/2)$$

所以第一项可以改写为：

$$\begin{aligned}&\int_0^{\nu\eta^2} \nu e^{-x/2}x^{\omega/2-1}2^{-\omega/2}/\Gamma(\omega/2)\mathrm{d}x\\&=\nu\times P(\chi_\omega^2<\nu\eta^2)\\&=\nu\times z\end{aligned} \qquad \cdots\cdots(\mathrm{B.7})$$

于是，对给定的自由度 ν，η 可以用上述方法计算，z 可以再用 GB/T 4086.2—1983 给出的 χ^2 分布表确定，从而(B.4)中右侧两项均可计算。

将式(B.2)、(B.5)、(B.6)和(B.7)代入式(B.4)，得：

$$\nu/\xi^2 = \nu\times z + 0.1\nu\eta^2$$

即

$$\xi = 1/\sqrt{z+0.1\eta^2} \qquad \cdots\cdots(\mathrm{B.8})$$

由此可确定 GB/T 6379 本部分表 23 所给出的修正系数 ξ 的值。

附 录 C
（资料性附录）
稳健分析中所用公式的推导

6.2.6 中用以计算平均值和标准差的稳健值的公式(62)和(63)，可根据 6.2.4 中算法 A 中的式(60)和式(61)导出。

使用 6.2.4 和 6.2.6 中的记号：

$$x^{*}=\sum x_{(i)}^{*}/p \quad\cdots\cdots(\text{C.1})$$

$$x'=\sum{}' x_{(i)}/(p-u_{\mathrm{L}}-u_{\mathrm{U}}) \quad\cdots\cdots(\text{C.2})$$

$$s'=\sqrt{\sum{}'(x_{(i)}-x')^{2}/(p-u_{\mathrm{L}}-u_{\mathrm{U}}-1)} \quad\cdots\cdots(\text{C.3})$$

其中 $\sum{}'$ 在满足 $|x_{(i)}-x_{(i)}^{*}|\leqslant\phi$ 的 $(p-u_{\mathrm{L}}-u_{\mathrm{U}})$ 数据项 $x_{(i)}$ 上求和。

因此式(C.1)可以写成：

$$\begin{aligned}p\times x^{*}&=\sum x_{(i)}^{*}\\&=\sum{}' x_{(i)}+u_{\mathrm{L}}\times(x^{*}-1.5s^{*})+u_{\mathrm{U}}\times(x^{*}+1.5s^{*})\end{aligned}$$

故

$$(p-u_{\mathrm{L}}-u_{\mathrm{U}})\times x^{*}=(p-u_{\mathrm{L}}-u_{\mathrm{U}})\times x'+(u_{\mathrm{L}}-u_{\mathrm{U}})\times 1.5s^{*}$$

从而得到推导式(C.4)过程中的

$$(p-u_{\mathrm{L}}-u_{\mathrm{U}})\times x^{*}=(p-u_{\mathrm{L}}-u_{\mathrm{U}})\times x'+(u_{\mathrm{L}}-u_{\mathrm{U}})\times 1.5s^{*} \quad\cdots\cdots(\text{C.4})$$

此即 6.2.6 中的式(62)。

为从式(61)推出式(63)，注意到式(61)中的求和项可以展开如下：

$$\sum(x_{(i)}^{*}-x^{*})^{2}=\sum{}'(x_{(i)}-x^{*})^{2}+(u_{\mathrm{L}}+u_{\mathrm{U}})\times(1.5s^{*})^{2} \quad\cdots\cdots(\text{C.5})$$

将式(C.4)代入此处的 x^{*}，经过一些代数运算得到：

$$\sum(x_{(i)}^{*}-x^{*})=\sum{}'(x_{(i)}-x')^{2}+(1.5s^{*})^{2}/(pu_{\mathrm{L}}+pu_{\mathrm{U}}-4u_{\mathrm{L}}u_{\mathrm{U}}) \quad\cdots\cdots(\text{C.6})$$

根据式(C.3)s'的定义，上式可以写成：

$$\begin{aligned}\sum(x_{(i)}^{*}-x^{*})&=(p-u_{\mathrm{L}}-u_{\mathrm{U}}-1)\times(s')^{2}\\&\quad+(1.5s^{*})^{2}(pu_{\mathrm{L}}+pu_{\mathrm{U}}-4u_{\mathrm{L}}u_{\mathrm{U}})/(p-u_{\mathrm{L}}-u_{\mathrm{U}})\end{aligned} \quad\cdots\cdots(\text{C.7})$$

将式(C.7)代入式(61)，即得式(63)。

附 录 D
（资料性附录）
参 考 文 献

[1] BS 812-103:1985, *Testing aggregates—Part 103:Methods for determination of partice size distribution*. British Standards Institution.

[2] BS 812-121:1989, *Testing aggregates—Part 121:Methods for determination of soundness*. British Standards Institution.

[3] BS 3144:1968, *Methods of sampling and physical testing of leather*. British Standards Institution.

[4] Scheffe, H.., *The analysis of variance*. Wiley, New York, 1959.

[5] Sweeney, An inter-laboratory study of the determination of protein by combustion in feeds. *Journal of the association of Official Analytical Chemists*, 72, 1989, pp. 770-774.

[6] ANALYTICAL METHODS COMMITTEE. Robust statistics—How not to reject outliers. Part1: Basic Concept. Part 2: inter-laboratory trials. *The Analyst*. 114, 1989, pp. 1653-1697 (part 1), pp. 1699-1702(part 2). Royal Society of Chemistry, London.

[7] Youden, W. J. The Youden plot. *Industrial Quality Control*, 15, 1959, pp. 133-137.

[8] Mandel, J. and Lashof, T. W. Interpretation and Generalization of Youden's Two-Sample Diagram. *Journal of Quality Technology*, 6, 1974, pp. 22-36.

[9] GB/T 4086.2—1983《统计分布数值表χ^2分布数值表》