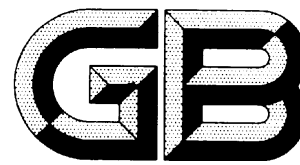


ICS 03.120.30
A 41



中华人民共和国国家标准

GB/T ××××-201×

实验室分析仪器测试方法 检出限和定量限的有效性确认评估

Standard Practice for Valid Verification of Detection and Quantitation
Estimate on Analytical Instrument Testing Method in Laboratory

(征求意见稿)

201×-××-××发布

201×-××-××实施

中华人民共和国国家质量监督检验检疫总局 发布
中国国家标准化管理委员会

目 次

前 言	2
1 范围	3
2 规范性引用文件	3
3 术语和定义、符号	4
3.1 术语和定义	4
3.2 符号	5
4 概要	5
5 使用意义	6
6 程序	6
6.1 IDE和IQE的计划、设计与方案	6
6.1.1 分析物、样品和方法的选择	6
6.1.2 IDE和IQE的设计选择	6
6.1.3 方案的选择	6
6.1.4 允许变异源的选择	6
6.2 IDE和IQE实施、数据筛选及模型选择	7
6.3 IDE和IQE的计算	9
6.3.1 IDE的计算	9
6.3.2 IQE的计算	10
6.4 删失数据的不可忽略量	10
7 报告	10
附录A（规范性附录）IDE和IQE的直线模型计算示例	11
A.1 ILSD模型的识别与拟合	11
A.2 模型R的拟合	12
A.3 LDE的递归迭代	12
A.4 LQE的计算	12
附录B（资料性附录）LDE和LQE的可选用模型	13
B.1 LDE的指数模型	13
B.1.1 指数模型的计算	13
B.1.2 指数模型的合理性评估	13
B.1.3 s_k 的计算	13
B.1.4 w_k 的计算	13
B.1.5 YC和LC的计算	13
B.1.6 LD的计算	13
B.2 LQE的混合模型	13
B.2.1 混合模型的选择	13
B.2.2 混合模型的合理性评估	13
B.2.3 s_k 的计算	14
B.2.4 s_k 的非线性最小二乘拟合	14
B.2.5 w_k 的计算	14
B.2.6 混合模型的计算	14
参考文献	15

前 言

本标准参考采用了以下两个 ASTM 标准：

ASTM D6091：2007 《校准误差忽略下分析方法 99%和 95%检出限的实验室间评估》（Standard Practice for 99%/95% Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error）。

ASTM D6512：2007 《实验室间定量限评估》（Standard Practice for Interlaboratory Quantitation Estimate）。

本标准的附录 A 为规范性附录。

本标准的附录 B 为资料性附录。

本标准由全国认证认可标准化技术委员会提出并归口。

本标准起草单位：

本标准主要起草人：

本标准于 201X年首次发布。

实验室分析仪器测试方法检出限和定量限的有效性确认评估

1 范围

本标准规定了校准误差忽略下实验室间 99%/95%检出限（以下简称 *IDE*）的应用评定指南，还规定了 *Z*%相对标准差的实验室间定量限（以下简称 *IQE*）的应用评定指南，并配套给出 *IDE* 和 *IQE* 的计算程序。

本标准规定了实验室间检出限估计中单个实验室在 90%容忍区间下的最小浓度计算，其检出概率为 95%，空白样品测量时未检出概率为 99%；还规定了实验室间定量限估计中单个实验室的最低浓度 *Z*%计算（符合文献[1]和文献[2]的定量方法），其中的 *Z* 通常取整数倍数（例如，10、20 或 30，也可小于 10）。

本标准规定了实验室间所用样品、浓度和程序的假定，为制订的检测方法适用范围提供了一个代表性的合理评定。本标准的正确实施，可确保 99%/95%的 *IDE* 和 *IQE* 能达到下述特性：

可获得日常的 *IDE* 和 *IQE*

——在日常分析的标准测量系统使用中，大部分实验室都能以合理的成本获得 *IDE* 和 *IQE* 性能，在实际情况下这种特性的需求也是可行的。使用代表性实验室的数据来计算 *IDE* 和 *IQE*。

——可涵盖日常的误差源

理想情况下 *IDE* 和 *IQE* 应包括偏倚源和测量过程的偶然变异，这些误差源涉及到但不限于：仪器固有的误差，某些特定的传递误差，以及实验室、操作者、样品制备和仪器之间的差异。

——可避免误差源的排除

理想情况下 *IDE* 和 *IQE* 应排除可避免的偏倚和变异源，即在日常现场测量中那些予以合理避免的误差源。这些误差源涉及到但不限于：对样品、测量步骤或确认方法中测量设备的变更，以及明显可辨别的抄写误差（假设会察觉且能修正或排除）。

——*IDE* 的错误检出概率低

IDE 是相应于某个被测浓度阈值（被测临界值）的一个真浓度，其给出很高的未检出的概率，即 99%（错误检出的概率， $\alpha=1\%$ ）。由此，当测量空白样品时，分析物未检出的概率应是 99%。为了保证使用性，需表明所用的特定样品，不能使用试剂空白。

——*IDE* 的错误未检出概率低

IDE 是一个真浓度，其给出很高的检出的概率，即 95%（错误未检出的概率， $\beta=5\%$ ）的同时，也给出错误检出的低概率。由此，当测量样品的 *IDE* 时，检出概率至少应是 95%。为了保证使用性，需表明所用的样品基质，不能使用试剂空白。

注：所指概率 α 和 β 是 *IDE* 风险评估的重要参数。

本标准适用于校准误差小于其它变异源的检测方法评价，这些主要变异源存在于以下之一：

——样品制备

所用的校准标准与样品制备不相关。

——操作者差异

操作者很少有机会会影响到校准结果（例如，使用自动校准）。

——实验室差异（任何理由）

识别与消除可能有困难。

——仪器（测量设备）差异

这些差异表现在：生产厂、型号、硬件、电子、抽样率、化学过程速率，积分时间（合成时间）（应该是色谱检测的“积分时间”）、软件计算、内部信号处理与门限值、有效样品容积、以及污染水平。

本标准适用于其它的数据质量目标，可选择其它 α 和 β 值来计算 *IDE*。然而，本标准仅给出 99% 和 95% 的 *IDE*。

本标准适用于 *IQE* 的数据质量目标，通常计算特定方法给出数据集下最低的相对标准差，若可能，应计算 *IQE*_{10%} 值。如果数据表明方法受噪声干扰加大，则不妨使用 *IQE*_{20%} 甚至是 *IQE*_{30%}。尽管相对标准差的水平小于 10% 的 *IQE* 可接受（例如 *IQE*_{1%}），但对具有较高相对标准差水平的 *IQE* 不予考虑（例如 *IQE*_{50%}）。合适的相对标准差水平依赖于 *IQE* 的使用目的。

本标准也适用于分析方法校准误差忽略下实验室间 99%/95% 检出限和定量限的评估。

2 规范性引用文件

下列文件中的条款通过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 XXXX 实验室质量控制 非标准测试方法的有效性评价 线性关系

GB/T 22554 基于标准样品的线性校准 (ISO 11095 MOD)

ASTM D7510 Standard Practice for Performing Detection and Quantitation Estimation and Data Assessment Utilizing DQCALC Software, based on ASTM Practices D6091 and D6512 of Committee D19 on Water

3 术语和定义、符号

下列术语和定义、符号适用于本标准。

3.1 术语和定义

3.1.1 实验室间 99%/95%的检出限评估 99%/95% Interlaboratory Detection Estimate, IDE

等同于 Currie 定义的 LD (见文献[1])。表示一个最低浓度, 在该浓度下有 90%的把握保证: 由胜任实验室参加的实验室间比对中, 一个实验室给出单测量的正确检出概率至少为 95%、以及空白样品浓度下正确未检出概率至少为 99%。

3.1.2 实验室间 $Z\%$ 定量估计 $Z\%$ Interlaboratory Quantitation Estimate, $IQE_{Z\%}$

等同于 Currie 定义的 LQ (见文献[1])。表示在该最低浓度下由胜任实验室参加的实验室间比对中, 一个实验室给出单测量有 $Z\%$ 的相对标准差 ($Z\%$ 的相对标准差计算来自实验室间标准差)。

3.1.3 删失测量 Censored Measurement

一种既不报告数值、也不报告丢失、而是报告“未检出”或“小于”(例如, “小于 0.1ppb”)的测量。不报告数值的原因有两个, 其一, 认为测量不具足够的精度或准确度; 其二, 怀疑分析物的存在。“小于”报告值的解释类似, 但还暗指(可能错误)任何大于或等于相关浓度值(例如, 0.1ppb)都可测量, 且能给出报告数值。

3.1.4 检出限 Detection Limit, LD

又称为检出的界限 LD 。是一种以物理单位或比例所表示的数值, 用以代表可信检出的最低水平。该水平以高概率区别于 0, 同时以高概率允许空白样品测量的未检出。

注: 某些情况下, 区别来自于非 0 值, 例如, 背景水平。同样注意, LD 也受测量和检出过程的日常误差源影响。 IDE 是 LD 的一个示例。

3.1.5 错误检出概率 Probability of False Detection

表示为 α 的错误正概率, 是一个空白样品的单测量检出 (见图 1)。这种概率通常称为第 I 类错误概率, 基于分析物、测量系统、分析方法、样品、操作者、以及用于确定检出是否出现的测量 (回收) 阈值 (测量临界值)。该定义通常指不想要的检出, 其来自于分析物非 0 浓度而不是空白的样品单测量, 假设这种非 0 浓度小于 LD 或 IDE 。

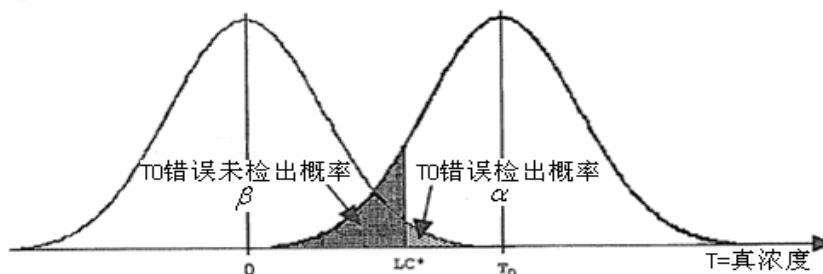


图 1 可信检出的简单示例

3.1.6 错误未检出概率 Probability of False Nondetection

表示为 β 或 $\beta(T)$ 的错误负概率, 是一个包含相关分析物非 0 浓度 T 下样品单测量的未检出, 是对真检出概率的补充 (见图 1)。这种概率函数通常称为第 II 类错误概率函数, 明显地基于浓度 (T)、以及相关因素: 分析物、测量系统、分析方法、样品、操作者和检出临界值。

3.1.7 正确检出概率 Probability of True Detection

表示为 $1 - \beta$ 或 $1 - \beta(T)$ 的概率, 是一个包含相关分析物非 0 浓度 T 下样品单测量的检出 (见图 1)。这种概率通常称为统计功效或检出功效, 明显地基于浓度 (T)、以及相关因素: 分析物、测量系统、分析方法、样品、操作者和检出临界值。

3.1.8 正确未检出概率 Probability of True Nondetection

表示为 $1 - \alpha$ 的正确负概率, 是一个未检出的空白样品的单测量, 是对错误检出概率的补充 (见图 1)。这种概率同样基于分析物、测量系统、分析方法、样品、操作者和响应阈值。正确未检出概率的描述同理: 适用于任何小于 DL 或 IDE 的非 0 浓度下样品单测量。

3.1.9 总体 $100(1 - \gamma)\%$ 统计容忍区间限 Confidence Statistical Tolerance Limit for $100(1 - \gamma)\%$ of a Population

又称为单侧统计容忍区间，是一个 $100(1-\gamma)\%$ 分位数的统计测定限，其在 $100(1-\gamma)\%$ 容忍区间内会超出（或落入）总体的 $100(1-\gamma)\%$ ，进一步解释和取值详见文献[2]。

3.1.10 定量限 Quantitation Limit, LQ

又称为定量的界限 LQ 。是一种以物理单位或比例所表示的数值，用以代表可信定量的最低水平。 IQE 是 LQ 的一个示例。

3.2 符号

T ——真浓度

T_k —— k 水平的 T 值

T_0 ——空白样的 T 值

模型 R ——实验室间测量浓度与 T 之间的平均回收直线模型

$ILSD$ ——实验室间标准差的模型

SD ——实验室间测量的样品标准差

OLS , WLS ——分别为普通和加权最小二乘

LC ——临界限，90%容忍区间下空白样测量值不超出1%的预期 T 回收值。

YC —— LC 测量期望值的估计，90%容忍区间下空白样测量值不超出1%。

IDE_0 , IQE_0 ——分别为 IDE 和 IQE 的初始估计值

g , h —— $ILSD$ 的拟合常数

error——误差，可为加性或乘性。

s_k , $\hat{s}_k$ —— k 水平下 T 的标准差及其估计值

$\hat{s}(0)$ —— $G(0)$ ，空白样品测量的标准差预测值

YD —— LD 测量期望值的估计

r_k , q_k ——残差值

a , b ——分别为模型 R 的截距和斜率估计值

$s_{\max}$, $T_{\max}$ ——分别为 T_k 和 $\hat{s}_k$ 的最大值

$f(g)$, $f(h)$ ——分别为 $g=g_0$ 和 $h=h_0$ 处的导函数

Q ——中间变量，用于 $ILSD$ 模型选择中曲线显著性的统计检验

p , u , v , c , d , p , q ——中间统计量

ss_k ——混合模型的迭代拟合值

4 概要

4.1 通过实验室间比对来评定一个检测方法最低浓度水平的可信检出。实验室间比对应该包含正确的不确定度模型 R （不建议外推）。同样，实验室间比对的设计和实施，应允许可以观察到特定影响水平下测量变异的日常已知源。实验室间比对实施后，可遵循一个可接受的科学程序，对离群实验室和单个测量进行剔除，每个浓度水平下所保留的数据应至少来自于6个独立实验室。

4.2 对所保留的数据进行分析，用来识别和拟合提出的3个 $ILSD$ 模型之一，这些模型描述了 $ILSD$ 与 T 之间的关系。识别过程包括对最简单模型到最复杂模型的依次评价， IDE 涉及到常数、直线和指数， IQE 涉及到常数、直线和混合（见文献[3]）。评价包括统计显著性和残差分析。

4.3 所选模型用于研究范围内所有浓度下 $ILSD$ 的预测，若 SD 不是常数，则需使用预测值的 WLS （否则 OLS ）进行模型 R 的拟合。使用残差分析来评价模型 R 的统计显著性和失拟误差。 $ILSD$ 还可用于空白测量的 SD 估计，使用该估计值来计算 YC 值。在胜任实验室为代表参加的实验室比对中， YC 值在90%容忍区间内不会超出所有空白测量的99%，并给出 LC 值。 $ILSD$ 模型还可预测浓度范围内的 SD 。在90%容忍区间内的 IDE 直接或迭代计算值，超出 YC 值的概率至少为95%，同时未超出空白样测量的概率为1%。

4.4 基于模型，通过直接或交互计算，利用 $RSD=10\%$ （ $Z=10$ ）估计值，求得最低浓度 $IQE_{10\%}$ 。若不存在这样的浓度，可以计算 $IQE_{20\%}$ 或 $IQE_{30\%}$ 来取代。如果能有数据质量目标支持， $IQE_{Z\%}$ 可计算到 $Z<10$ 的数值。

5 使用意义

5.1 本标准的正确使用，有助于大多数实验室对检测方法研究下的 IDE 和 IQE 获取，即大多数实验室都能对大于 $IQE_{Z\%}$ 的浓度进行测量。对于相同样品同一分析物的低水平浓度下可信检出， IDE 和 IQE 都为

胜任实验室的检测方法预期使用提供了依据。

5.2 IDE 用于相同样品的同一分析物下不同方法检出功效的分析比较。**IQE** 用于相同样品的同一分析物下不同方法定量能力的分析比较。

5.3 IDE 给出 95% 的高概率, 所用方法结果值超出 **IDE** 表示样品中存在分析物, 而 **IDE** 的 99% 高概率表示空白样品未检出。

5.4 数据使用的 **IDE** 和 **IQE** 很重要, 由此本标准建立了检测方法应用下的实验室间检出和定量能力。

6 程序

6.1 **IDE** 和 **IQE** 的计划、设计与方案

6.1.1 分析物、样品和方法的选择

至少选择一个相关的分析物, 通常选择的是痕量或接近痕量浓度水平的分析物, 例如, 被控制和规定的毒性物质。针对每个分析物, 根据以下考虑来选择最大的 **T**。

6.1.1.1 选择的最大 **T** 应超出 **IDE** 或 **IQE** 的预期值 2 倍或更多。

6.1.1.2 模型 **R** 应涵盖自 0 到所选最大 **T** 的整个范围。

6.1.1.3 模型 **R** 应描述自 0 到最大 **T** 范围内的 **SD**。

6.1.1.4 应具有足够的范围以便估计 **ILSD** 和模型 **R** 在统计上系数的显著性。也可选择一个或多个样品, 采用一个可接受的标准分析方法对分析物进行研究。如果不存在可能的样品干扰, 才有必要测定一系列的可接受样品, 而非选择特定的样品。

6.1.2 **IDE** 和 **IQE** 的设计选择

应根据 **ILSD** 的预先模型和有关标准来设计。**ILSD** 的预先模型有助于 **IDE** 和 **IQE** 研究的选择。**IDE** 研究建议选择: 常数模型、直线模型和指数模型; **IQE** 研究建议选择: 常数模型、直线模型和混合模型。化学、物理、实验证据、或信息判断有助于其中的模型选择, 但不可能预测 **SD** 与 **T** 之间的关系。

6.1.2.1 **IDE** 研究的 **T** 至少选择 5 个, 假定设计包含一个空白、至少一个近似 $2 \times IDE_0$ 、以及至少一个低于 IDE_0 的非 0 浓度。**IQE** 研究的 **T** 至少选择 7 个, 假定设计包含至少一个近似 $2 \times IQE_0$ 、至少一个低于 IQE_0 的非 0 浓度、以及一个空白或未添加样品。

6.1.2.2 实际上, 所研究的浓度水平 **T** 应为已知或可知, 可订购或得到的标准样品, 其浓度视为已知; 可获得准确度的测定 (例如, 多个实验室的中位数、或来自于使用较高准确度方法的胜任实验室结果), 其浓度视为可知。

6.1.3 方案的选择

包括设计的安排次序与系统需调整的细节、以及使用额外空白样操作, 等等, 还要考虑可能出现的问题, 例如, 传送、研究成本 (时间和金钱)、测量系统漂移的时间常数、以及样品的降级。

6.1.3.1 实验室间比对的监督者应通知所参加的实验室, 去掉 (若可能) 所有内部测量系统的阈值 (比如, 仪器的检出限或峰面积的阈值), 因为这些阈值被用来确定是否作为 “未检出”、“小于”、或 “删失” 的测量数值报出。如果 “删失” 不可避免, 则应报告实验室研究数据的 “删失” 阈值。

6.1.4 允许变异源的选择

假定在任何 **T** 下, 实验室间测量所汇集的多种变异源处于正态分布。提供可用数据的实验室数目应尽可能多, 以便能发现代表性的实验室间变异。应允许通常的实验室内变异反映出日常测量过程中的情况。理想情况下, 应有多个参加实验室, 每个实验室由不同操作者使用不同 (合格) 的测量系统, 在不同天数以随机形式来实施一个可信的盲样测量。

6.1.4.1 参加实验室个数的尽可能多是确保研究成功的重中之重, 其个数要超出提交整套可用数据的次数。建议参加实验室的个数为 10。

6.1.4.2 研究的实施应尽可能模拟日常的实验室测量, 尤其是不同于自动程度高的那些人工操作方法。即说, 操作者不应得知这些样品的真实浓度, 也不能得知其正在测量的特定研究样品。这些限制会最大限度减少分析研究中经常过于关注而导致数据不真的风险。

6.1.4.3 由于允许不同操作者实施方法的变异存在, 且这些变异在日常的分析中可观察到, 则每个实验室应尽可能安排多位胜任的操作者来参加研究。

6.1.4.4 由于在设备和维护中存在模型间和仪器间的差异, 且这些差异在日常的分析中可观察到, 则每个实验室应尽可能使用多个良好的测量系统。

6.1.4.5 由于分析实验室环境条件下天数间变化, 污染、溶剂纯度及其它因素都会影响到测量, 且这些影响因素在日常的分析中可观察到, 则每个实验室应尽可能规定出所需的天数, 妥善安排研究。

6.2 IDE 和 IQE 实施、数据筛选及模型选择

6.2.1 实验室无需实施空白修正，除非检测方法要求这种扣减。每个实验室应提交方法空白以及未经修正的测量数据，研究的监督者可决定所报告的测量是否予以修正。

6.2.2 对所保留的数据进行筛选。在任一浓度下，若所保留数据有 10% 属于“非检出”或“小于”，进入 6.4。

6.2.3 ILSD 模型的识别与拟合——按以下步骤来识别 ILSD 模型和估计模型的系数。

文献[4]、文献[5]和文献[6]详细讨论了分析化学中 SD 建模和 WLS 做法。ILSD 模型描述实验室间实际测量的 SD 与 T 之间的未知（或部分已知）关系 $[s=G(T)]$ ，该模型用于两个目的：对模型 R 进行 WLS 回归拟合；提供实验室间所需 SD 估计值以便确认 LC、IDE 和 IQE。

6.2.3.1 建议使用 2 个 ILSD 模型（IDE 的指数模型和 IQE 的混合模型选择见附录 B）。依次拟合常数模型和直线模型，直至得到适宜的模型。若能引用适当的参考文献，也可结合先验信息和实验结果，依次考虑模型的选择。

a) 常数模型：

$$s=g+\text{error} \dots \dots \dots (1)$$

式中：

s ——SD

b) 直线模型：

$$s=g+h \times T+\text{error} \dots \dots \dots (2)$$

在所有情况下都假定 $g>0$ （这种限定与附录 B 的混合模型不相关）， $g<0$ 没有实际意义，表明应使用不同的 ILSD 模型。另外假定，由于空白或其它低浓度样品中的数据删失， g 不可低估。

如果 $h<0$ ，表明统计上小于 0，此时应评定常数模型。

6.2.3.2 ILSD 模型的识别与拟合程序

a) 离群数据剔除后，合并所有的保留数据。

b) 针对每个 T_k ，计算 ILSD 调整后的实验室间样品标准差 s_k ，视为实验室间测量的真实估计值 σ_k 。 s_k 值由表 1 中偏倚修正因子 a'_n 相乘求得。本标准中所有涉及到的实验室间样品标准差的计算和拟合值均视为 s_k 。

表 1 特定浓度下测量样品标准差的调整因子偏倚修正

n	2	3	4	5	6	7	8	9	10
a'_n	1.253	1.128	1.085	1.064	1.051	1.042	1.036	1.031	1.028

针对每个 T_k ，应建立调整值 $s_k=a'_n s'_k$ 的模型，以取代样品标准差 s'_k 。当 $n>10$ 时，使用式 $a'_n=1+[4(n-1)]^{-1}$ ，见文献[7]和文献[8]。

c) 基于 T_k ，对 s_k 作图。

d) 使用 OLS，针对 T_k 进行 s_k 的回归，暂时假定直线模型有效，其关系式中给出系数 g 和 h ：

$$s_k=g+h \times T_k+\text{error} \dots \dots \dots (3)$$

e) 常数模型的合理性评价需做两项检查：首先，检查 OLS 回归斜率估计值 h 的 p 值，若 $p<0.05$ （可参见 GB/T XXXX 的 t 检验），则统计表明斜率显著，应拒绝常数模型。其次，检查上述 c)、或 OLS 拟合的残差作图，如果存在明显的趋势（例如，二次或指数类型特征），应拒绝常数模型，进行步骤 h)。若无法拒绝常数模型，进入模型 R 的拟合。也可在 s_k 对 T_k 的作图分析中，若 s_k 呈水平独立性，且由专家组研究决定后，可采用常数模型评价。

f) 由于统计上斜率显著，拒绝常数模型，按下式计算残差值，并基于 T_k 对 s_k 作图。

$$r_k=s_k-(g+h \times T_k) \dots \dots \dots (4)$$

g) 评价直线模型的合理性。检查 (f) 的作图。如果在浓度范围内发现明显的趋势（例如，二次或指数类型特征），应拒绝直线模型（若需做进一步的评估，IDE 要考虑指数模型的拟合，而 IQE 要考虑混合模型的拟合，详见附录 B）。若无法拒绝直线模型，进入模型 R 的拟合。也可在 s_k 对 T_k 的作图分析中，若 s_k 呈水平依赖性，且由专家组研究决定后拒绝常数模型。然后检查 r_k 对 T_k 的 OLS 作图，如果未出现明显的趋势，应采用直线模型评价。

6.2.4 模型 R 的拟合——模型 R 为一条简单的直线：

$$Y=a+bT+\text{error} \dots \dots \dots (5)$$

拟合步骤来自于 6.2.3 的模型选择，若选择常数模型，可用 OLS 来拟合模型 R （见表 2）。若选择直线模型，则可用 WLS 来拟合模型 R ，这种拟合近似给出系数 a 和 b 的最小方差线性无偏估计。

WLS 近似给出系数的最小方差无偏线性估计值 a 和 b （见表 2）。

表2 模型 R 系数估计的 OLS 和 WLS 计算

OLS	WLS
$\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i$	$\bar{T}_w = \frac{\sum_{i=1}^n w_i T_i}{\sum_{i=1}^n w_i}$
$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$	$\bar{Y}_w = \frac{\sum_{i=1}^n w_i Y_i}{\sum_{i=1}^n w_i}$
$S_{TT} = \sum_{i=1}^n (T_i - \bar{T})^2$	$S_{wTT} = \sum_{i=1}^n w_i (T_i - \bar{T})^2$
$S_{TY} = \sum_{i=1}^n (T_i - \bar{T})(Y_i - \bar{Y})$	$S_{wTY} = \sum_{i=1}^n w_i (T_i - \bar{T})(Y_i - \bar{Y})$
斜率 $= b = S_{TY}/S_{TT}$	斜率 $= b = S_{wTY}/S_{wTT}$
截距 $= a = \bar{Y} - b\bar{T}$	截距 $= a = \bar{Y}_w - b\bar{T}_w$

6.2.4.1 直线模型的 WLS 步骤

a) 使用直线模型和系数估计值, 计算每个 T_k 下实验室间标准差的预测值 $\hat{s}_k$:

$$\hat{s}_k = g + hT_k \dots\dots\dots (6)$$

b) 计算 WLS 的加权值:

$$w_k = \frac{1}{(\hat{s}_k)^2} \dots\dots\dots (7)$$

d) WLS 的计算类似 OLS, 见表 2。模型 R 的系数估计结果为 a 和 b 。

g) 模型 R 拟合后应评价其合理性和失拟误差, 以确保:

- 拟合统计上显著 ($p < 0.05$);
- 失拟误差统计上不显著 ($p > 0.05$) (可参见 GB/T 22554);
- 以及残差作图无明显的趋势 (例如, 类似二次或指数特征)。

如果模型 R 的评价失败, 实验室间比对的监督者应决定是否仅用子集数据来进行分析 (可能模型的失败来自于较高 T 的拟合), 或是否需要提供更多的数据。

6.3 IDE 和 IQE 的计算

6.3.1 IDE 的计算

IDE 的计算是利用 ILSD 模型, 来估计 $T=0$ 和 IDE 下的实验室间标准差, 并使用模型 R 将被测浓度转换成 T , 反之亦然。计算分为三步, 依次的计算为: YC、LC 和 IDE, 也可计算 IDE 预期测量的 YD。

6.3.1.1 YC 的计算如下:

$$YC = k1 \times \hat{s} + a \dots\dots\dots (8)$$

式中:

$k1$ ——基于 n 观测 (见表 3) 的 90%统计容忍单侧上限 (也称为正态分布 99%分位数的单侧统计容忍区间)。

n ——IDE 研究后所保留的测量总数。

表3 90%容忍区间内 99%和 95%上限的 IDE 容忍因子

所保留观测值数, n	99%分位数, $k1$	95%分位数, $k2$
5	4.67	3.40
10	3.53	2.57
15	3.21	2.33
20	3.05	2.21
25	2.95	2.13
30	2.88	2.08
35	2.83	2.04
40	2.79	2.01
45	2.76	1.99
50	2.74	1.97
55	2.71	1.95
60	2.69	1.93
65	2.68	1.92
70	2.66	1.91
75	2.65	1.90
80	2.64	1.89
90	2.62	1.87
100	2.60	1.86
150	2.55	1.82
200	2.51	1.79

6.3.1.2 LC 的计算如下:

$$LC = \frac{YC - a}{b} \dots\dots\dots (9)$$

6.3.1.3 若使用常数模型，计算如下：

$$LD = LC + k2 \times \frac{\hat{s}(0)}{b} \dots\dots\dots (10)$$

式中：

$k2$ ——基于 n 观测（见表 3）的 90%统计容忍单侧上限（也称为正态分布 95%分位数的单侧统计容差区间）。

6.3.1.4 若使用直线模型， LD 的计算可采用递归处理：

$$LD_{i+1} = [k1 \times \hat{s}(0) + k2 \times \frac{(g + h \times T_k)}{b}] \dots\dots\dots (11)$$

LD_0 估计如下：

$$LD_0 = 2 \times LC \dots\dots\dots (12)$$

$$LD_0 = LC + k2 \times \frac{\hat{s}(0)}{b} \dots\dots\dots (13)$$

将每一次迭代所估计的 LD 值纳入到上述的递归式中，给出新的 LD 值，直至迭代给出连续 LD 估计值之差小于 1%。此时 LD 即为 T 值，其样品测量有 95%概率给出超过 YC 值。

6.3.1.5 结果有 $IDE = LD$ ，在 IDE 真浓度下单样品测量值有 95%概率超出 YC 值（报告检出），与此同时，单空白样测量仅有 1%概率超出 YC 值，两者均处于 90%容忍区间内。在 LD 真浓度下的样品测量预期值计算为：

$$YD = a + b \times LD \dots\dots\dots (14)$$

6.3.2 IQE 的计算

IQE 的计算是使用 $ILSD$ 模型来估计实验室间标准差、以及使用模型 R 来标度标准差。 IQE 有效计算值应处于研究中的所用浓度范围内，其通用计算式如下：

$$T = \frac{100}{Z} \times G(T) \dots\dots\dots (15)$$

式中：

$G(T)$ —— T 的实验室间标准差（以浓度单位）估计函数。

Z ——按升序取值 10、20 或 30。

首先求得 $LQ_{10\%}$ 计算值。如果 $LQ_{10\%}$ 不存在或超出所用浓度范围时，若可能需计算 $LQ_{20\%}$ 。如果 $LQ_{20\%}$ 不存在或超出所用浓度范围时，若可能则计算 $LQ_{30\%}$ 。如果特定使用适宜时， $IQE_{Z\%}$ 可计算为 $Z < 10$ 的任何值，但不建议 $Z > 30$ 。这样， IQE 的计算是基于函数 G 部分的 $ILSD$ 模型。

$Z' = \frac{100 \times h}{b}$ 可获得到 $LQ_{Z\%}$ ，例如， $Z' = \frac{100 \times 0.17}{1.0} = 17$ ，精确的 IQE 视为 $LQ_{20\%}$ （根据最接近的较高倍数 10）。

6.3.2.1 常数模型的计算有： $\hat{s} = g, G(T) = \frac{g}{b}, T = \frac{100}{Z} \times \frac{g}{b}$ ，则：

$$LQE_{Z\%} = \frac{100}{Z} \times \frac{g}{b} \dots\dots\dots (16)$$

6.3.2.2 直线模型的计算有： $\hat{s} = g + hT, G(T) = \frac{(g + hT)}{b}, T = \frac{100}{Z} \times \frac{(g + hT)}{b}$ ，则：

$$LQE_{Z\%} = \frac{g}{b \times \frac{Z}{100} - h} \dots\dots\dots (17)$$

6.4 删失数据的不可忽略量

至少一个 T 下存在大于 10%的数据报告为“未检出”或“小于”。尽管努力减少和排除了“未检出”或“小于”的报告值，但其仍会在水平下出现，影响到 6.2 和 6.3 的数据分析。如果存在过多的删失数据，实验室间比对研究的监督者应就此通知实验室，能否从测量数据中选取“未删失”的数据，若无法做到有效补救，就应认真考虑在 IDE 研究中增加新的和不同 T （通常较高 T ）样品测量。

7 报告

7.1 报告内容

a) 人员和分析方法的识别、分析物、基体、样品特性、以及所用特定方法的选择。

- b) 研究中发现的任何异常现象。
- c) 数据筛选程序的描述与所保留的结果、离群值的剔除、丢失值、以及与最初数据组相比的所用数据百分比。
- d) 数据的统计分析与 *LC*、*IDE* 和 *IQE* 计算。
- e) 所选的 *ILSD* 模型，并给出理由。
- f) *ILSD* 模型和模型 *R* 的系数估计值。

7.2 确认内容

- a) 数据的誊抄与报告是否正确无误。
- b) 数据的分析是否正确无误。
- c) 分析结果使用的适宜性，包括计算所需假设的可能拒绝。
- d) 整个最终报告的完整性评估，作为 *IDE* 和 *IQE* 测定的文件化内容。

7.3 审查内容

- a) 报告中是否附有审查与结果的表述。
- b) 审查人员是否胜任应用统计方面的能力。
- c) 审查人员是否胜任分析化学方面的能力。

A.1 *ILSD*模型的识别与拟合

A.1.1 表 A.1 给出了 $n=10$ 的 T 水平下提交的测量结果、包括空白样的 7 个 T 水平 ($T_k=0、0.5、1.0、2.0、4.0、8.0$ 和 $12.0\text{ }\mu\text{g/L}$)、以及 T_k 水平下的 s_k 计算值。本示例假定，数据未被剔除，且遵循本标准规定的程序进行。

表 A.1 IDE 和 IQE 测量报告与统计计算

T_k $\mu\text{g/L}$	实验室内比对试验结果, Y_i $\mu\text{g/L}$	S_k	$\hat{s}_k$	r_k	W_k
0.0	0.105, 0.263, 0.293, 0.187, 0.106, 0.329, 0.080, 0.524, 0.287, 0.206	0.1729	0.0649	0.1078	237.18
0.5	0.354, 0.724, 0.682, 0.327, 0.527, 0.868, 0.730, 0.434, 0.794, 0.642	0.1929	0.1283	0.0647	60.74
1.0	1.241, 0.668, 1.200, 1.370, 1.106, 0.964, 0.949, 1.421, 1.032, 1.134	0.2270	0.1917	0.0353	27.22
2.0	2.174, 2.388, 2.153, 2.366, 2.306, 2.309, 1.663, 2.841, 1.933, 1.809	0.3449	0.3184	0.0262	9.86
4.0	3.660, 3.734, 3.167, 3.578, 4.278, 3.383, 3.873, 4.479, 3.919, 3.856	0.3995	0.5719	-0.1725	3.06
8.0	6.592, 7.520, 6.822, 7.751, 7.771, 7.296, 8.578, 6.863, 7.840, 8.821	0.7521	1.0790	-0.3270	0.86
12.0	9.496, 9.081, 13.942, 10.547, 9.324, 13.148, 10.994, 11.774, 12.320, 13.521	1.8519	1.5860	0.2655	0.40

A.1.2 表 A.1 中的结果作图见图 A.1。使用 *OLS* 来对 s_k 进行直线模型拟合，表明 s_k 随 T 水平而增进，建议采用直线模型。

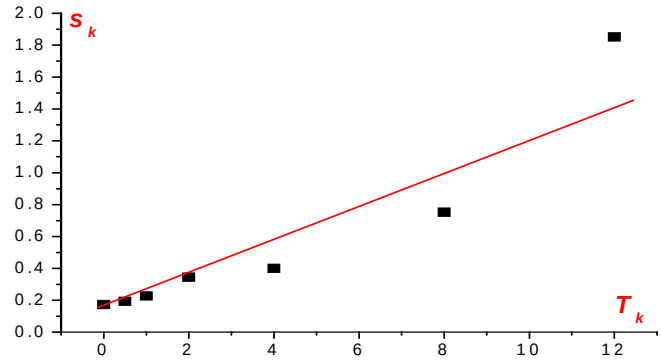


图 A.1 T 水平下的 s_k 拟合

A.1.3 所选直线模型的残差拟合值计算如下：

$$r_k = s_k - (0.0649 + 0.1268 \times T_k) \dots\dots\dots (\text{A.1})$$

A.1.4 表 A.1 中的结果作图见图 A.2。基于 T_k 水平的 r_k 作图表明不存在明显的趋势，且由专家组研究决定后，应采用直线模型评价。

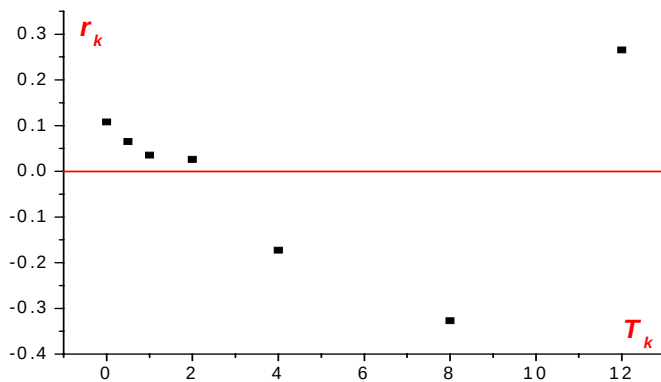


图 A.2 T 水平下的 r_k 拟合

A.1.5 直线模型的截距和斜率估计值分别是 $g=0.0649$, $h=0.1268$ 。检查 *OLS* 回归斜率估计值 h 的 p 值， $p<0.05$ （可参见 GB/T XX 的 t 检验），统计表明斜率显著，应拒绝常数模型。鉴于 A.1.2 的作图中 s_k 呈水平依赖性、以及 A.1.4 的残差作图中未出现明显的趋势（例如，二次或指数类型特征），则接受直线模型。

A.1.6 可以进行模型 R 的拟合。

A.2 模型 R 的拟合

A.2.1 因 T_k 水平下的标准差非呈常数，则使用 WLS 来拟合模型 R 。

A.2.2 因直线模型为 $\hat{s}_k = g + h \times T_k$ ，则截距 g 和斜率 h 的估计给出每个 T_k 水平下的预测值 $\hat{s}_k$ （见表 A.1）。

A.2.2 计算 w_k 值（结果见表 A.1），按下式给出模型 R 的系数 $a=0.2042$ 、 $b=0.9228$ ：

$$\text{模型 } R: Y=a+b \times T+\text{error} \dots\dots\dots (A.2)$$

A.2.3 针对模型 R 的 a 和 b 进行统计检验，检验结果有：拟合的统计 $p<0.05$ 、失拟误差的统计 $p>0.05$ （可参见 GB/T 22554），残差作图无明显的趋势（例如，类似二次或指数特征）。接受模型 R 的拟合。

A.3 IDE 的递归迭代

A.3.1 按下式计算 YC ，其中： $k_1=2.64$ ， $g=0.0649$ ， $a=0.2042$ 。

$$YC = k_1 \times \hat{s}(0) + a = 2.69 \times 0.0649 + 0.2042 = 0.3788 \dots\dots\dots (A.3)$$

A.3.2 按下式计算 LC ，其中： $b=0.9228$ 。

$$LC = (YC - a) / b = (0.3788 - 0.2042) / 0.9228 = 0.189 \dots\dots\dots (A.4)$$

A.3.3 进行 LD 的递归计算（其中 $k_2=1.93$ ），初始值计算如下：

$$LD_0 = LC + k_2 \times \hat{s}(0) / b = 0.189 + 1.93 \times 0.0649 / 0.9228 = 0.325 \dots\dots\dots (A.5)$$

A.3.4 由此递归函数的迭代处理如下：

$$LD_1 = LC + k_2 \times (g + h \times LD_0) / b = 0.189 + 1.93 \times (0.0649 + 0.1268 \times 0.325) / 0.9228 = 0.411 \dots\dots\dots (A.6)$$

$$LD_2 = 0.189 + 1.93 \times (0.0649 + 0.1268 \times 0.411) / 0.9228 = 0.433 \dots\dots\dots (A.7)$$

直至收敛于第 8 次迭代， $LD_7 \approx LD_8 = 0.422 > (2 \times LC) = 0.378$ 。

A.3.5 YD 值的计算如下：

$$YD = a + b \times LD = 0.2042 + 0.9228 \times 0.442 = 0.612 \dots\dots\dots (A.8)$$

A.4 IQE 的计算

A.4.1 已知 $g=0.0649$ ， $h=0.1268$ ， $b=0.9228$ ，计算 $Z' = \frac{100 \times h}{b}$ ，近似获得 Z 值为 20。

A.4.2 IQE 的计算有：

$$LQE_{20\%} = LQE_{Z\%} = \frac{g}{b \times \frac{Z}{100} - h} = 1.123 \dots\dots\dots (A.9)$$

B.1 IDE 的指数模型

B.1.1 指数模型的计算

IDE 的指数模型计算公式如下：

$$s = g \times e^{hT} + error \dots\dots\dots (B.1)$$

$$s = g \times e^{hT} \times error \dots\dots\dots (B.2)$$

B.1.2 指数模型的合理性评估

首先检查 OLS 回归估计 h 的 p 值，若 $p < 0.05$ （可参见 GB/T XXXX 的 t 检验），表明统计上的斜率显著，则拒绝常数模型。然后，检查 r_k 对 T_k 的 OLS 作图，如果出现明显的趋势图形，则应拒绝直线模型。

为了评估指数模型的合理性，应对模型进行拟合，最简单方法是采用 OLS 的标准差对数回归：

$$\ln s_k = \ln g + hT_k + error \dots\dots\dots (B.3)$$

拟合应满足两种检查， h 的 p 值应小于 0.05、以及对数形式绘制的残差作图中不能出现趋势图形。在 r_k 对 T_k 作图中， r_k 值计算如下：

$$r_k = \ln s_k - (\ln g + hT_k) \dots\dots\dots (B.4)$$

若能满足上述两种检查，接受指数模型，进入模型 R 的拟合。

B.1.3 $\hat{s}_k$ 的计算

因 T 水平下的标准差非呈常数，使用 WLS 来拟合模型 R ，这种拟合近似给出系数 a 和 b 的最小方差线性无偏估计。使用 ILSD 模型和系数估计值，计算每个 T_k 下的预测值 $\hat{s}_k$ ：

$$\hat{s}_k = g \times e^{hT_k} \dots\dots\dots (B.5)$$

B.1.4 w_k 的计算

w_k 的计算见式 7，WLS 的计算见表 2。

B.1.5 YC 和 LC 的计算

同 6.3.1.1 和 6.3.1.2。

B.1.6 LD 的计算

同 6.3.1.4，其中，指数模型的递归计算有：

$$LD_{i+1} = [k1 \times \hat{s}(0) + k2 \times \frac{(g \times e^{h \times T_k})}{b}] \dots\dots\dots (B.6)$$

B.2 IQE 的混合模型

B.2.1 混合模型的选择

如果 IQE 的直线模型选择不妥，可选用混合模型，混合模型计算公式如下：

$$s = \sqrt{g^2 + (hT)^2} + error \dots\dots\dots (B.7)$$

B.2.2 混合模型的合理性评估

使用 OLS，基于 T_k 对 T_k^2 进行回归，产生拟合系数 u 和 v ，用于残差值 q_k 的计算：

$$q_k = (u + vT_k) - T_k^2 \dots\dots\dots (B.8)$$

使用 OLS，同时基于 T_k 和 q_k ，对 s_k 进行回归，产生拟合系数 g 和 h ，但增加了 Q 项：

$$s_k = g + hT_k + Qq_k + error \dots\dots\dots (B.9)$$

首先使用非线性最小二乘进行拟合，然后对混合模型的拟合做评估，以对数形式给出 r_k ，基于 T_k 作图， r_k 计算如下：

$$r_k = \ln s_k - \ln \hat{s}_k \dots\dots\dots (B.10)$$

如果作图表明不存在趋势图形，且检验 Q 的 p_Q 值有 $p_Q < 0.05$ 和 $Q > 0$ ，即可使用混合模型，进入

模型 R 的拟合。

B.2.3 $\hat{s}_k$ 的计算

使用 WLS 来拟合模型 R , 近似给出系数 a 和 b 的最小方差线性无偏估计, 计算每个 T_k 下的预测值 $\hat{s}_k$:

$$\text{混合模型: } \hat{s}_k = \sqrt{g^2 + (hT_k)^2} \dots\dots\dots (\text{B.11})$$

B.2.4 $\hat{s}_k$ 的非线性最小二乘拟合

a) 计算 T 水平下的 $\ln s_k$ 值, 令 j 为迭代次数, 设定 $j=0$, 则 g_0 和 h_0 有:

$$g_0 = s_1 \dots\dots\dots (\text{B.12})$$

$$h_0 = (s_{\max} - s_1) / (T_{\max} - T_1) \dots\dots\dots (\text{B.13})$$

b) 使用当前的 g_j 和 h_j 值, 计算混合模型的迭代拟合值 ss_k 和 $\ln ss_k$, 并计算 $r_k = \ln s_k - \ln ss_k$ 。

c) 分别计算 fg_k 和 fh_k 、以及中间统计量 u 、 v 、 c 、 d 、 p 和 q :

$$fg_k = g_j / e^{2\ln ss_k} \dots\dots\dots (\text{B.14})$$

$$fh_k = h_j (T_k)^2 / e^{2\ln ss_k} \dots\dots\dots (\text{B.15})$$

$$u = \sum_k (fg_k)^2 \dots\dots\dots (\text{B.16})$$

$$v = \sum_k (fh_k)^2 \dots\dots\dots (\text{B.17})$$

$$c = \sum_k (fg_k \times fh_k) \dots\dots\dots (\text{B.18})$$

$$d = 1 / (uv - c^2) \dots\dots\dots (\text{B.19})$$

$$p = \sum_k (fg_k \times r_k) \dots\dots\dots (\text{B.20})$$

$$q = \sum_k (fh_k \times r_k) \dots\dots\dots (\text{B.21})$$

d) 计算 g 和 h 的第 j 次变化:

$$\Delta g = d(vp - cq), \quad dg\% = 100 \left| \Delta g / g_j \right| \dots\dots\dots (\text{B.22})$$

$$\Delta h = d(uq - cp), \quad dhT\% = 100 \left| \Delta h / h_j \right| T_{\max} \dots\dots\dots (\text{B.23})$$

e) 计算新的 g 和 h :

$$g_{j+1} = g_j + \Delta g \dots\dots\dots (\text{B.24})$$

$$h_{j+1} = h_j + \Delta h \dots\dots\dots (\text{B.25})$$

f) j 的每 1 次迭代依此类推, 直至给出 $dg\% < 1\%$ 和 $dhT\% < 1\%$ 。

B.2.5 w_k 的计算

w_k 的计算见式 7, WLS 的计算见表 2。

B.2.6 混合模型的计算

因有 $\hat{s} = \sqrt{g^2 + (hT)^2}$, $G(T) = \frac{\sqrt{g^2 + (hT)^2}}{b}$, $T = \frac{100}{\frac{Z}{b}} \times \sqrt{g^2 + (hT)^2}$, 则:

$$LQ_{Z\%} = \frac{g}{\sqrt{\left(\frac{b \times Z}{100}\right)^2 - h^2}} \dots\dots\dots (\text{B.26})$$

- [1] L.A.Currie, ed., *Detection in Analytical Chemistry—Importance Theory, and Practice*, American Chemical Society, Chapter 1, 1988, p.10.
- [2] Hahn, G., and Meeker, W., *Statistical Intervals—A Guide for Practitioners*, Wiley, New York, 1991, pp. 34-36,41-52.
- [3] Rocke, D. M., and Lorenzato, S., "A Two-Component Model for Measurement Error in Analytical Chemistry," *Technometrics*, Vol 37, No.2, May 1995, pp. 176-184.
- [4] Caulcutt, R., and Boddy, R., *Statistics for Analytical Chemists*, Chapman and Hall, New York, 1983, pp. 79-85, 100-110, 114-123.
- [5] Draper, N., and Smith, J., *Applied Regression Analysis, 2nd Edition*, Wiley, New York, 1981, pp. 33-40.
- [6] Carroll, R., and Ruppert, D., *Transformations and Weighting in Regression*, Chapman and Hall, New York, NY, 1988, pp. 10-11, 62-69.
- [7] Maddalone, R., Rice, J., et al, "Defining Detection and Quantitation Limits," *Water Environment and Technology*, Vol 5, No.3, January 1993, pp.41-44.
- [8] Johnson, N., and Kotz, S., *Continuous Univariate Distribution*, Vol 1, John Wiley & Sons, Inc., New York, NY, 1970, p. 63.
-
-

中华人民共和国国家标准

GB/T XXXX—201X

实验室分析仪器测试方法

检出限和定量限的有效性确认评估

Standard Practice for Valid Verification of Detection and Quantitation
Estimate on Analytical Instrument Testing Method in Laboratory

编制说明

(征求意见稿)

国家标准编制组

二〇一〇年九月

GB/T XXXX—201X

实验室分析仪器测试方法检出限和定量限的有效性确认评估 Standard Practice for Valid Verification of Detection and Quantitation Estimate on Analytical Instrument Testing Method in Laboratory

编制说明

一、任务来源

本标准项目的制订，已列入国家质量监督检验检疫总局《2007 年制、修订国家标准项目计划》，项目序号为“国家质量监督检验检疫总局 20075324-T-469”。

本项任务由辽宁出入境检验检疫局提出、全国认证认可标准化技术委员会（TC261）归口、中国合格评定国家认可委员会负责，定于 2010 年底完成。

本项任务由辽宁出入境检验检疫局牵头、中国合格评定国家认可委员会（CNAS）、山东出入境检验检疫局等单位合作组成标准起草工作组，共同对标准的征求意见稿进行草拟整理工作。

本标准主要起草单位：

本标准主要起草人：

二、标准制订依据

本标准参考采用了以下 2 个 ASTM 标准：

ASTM D6091: 2007《校准误差忽略下分析方法 99%和 95%检出限的实验室间评估》（Standard Practice for 99%/95% Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error）。

ASTM D6512: 2007《实验室间定量限评估》（Standard Practice for Interlaboratory Quantitation Estimate）。

三、标准制订的目的和原则

适宜的分析方法会具有检出限和定量限，以满足化学潜在的风险评估要求。方法的检出限和定量限直接影响到数据的使用，尤其是靠近检出限的数据报告很有可能会给出 α 和 β 错误率。风险评估者必须争取化学工作者的帮助，来选择一個检出限接近所需行动水平的分析方法。如果检出限高于相关的浓度水平，则不要使用这个分析方法。

本标准假定 x 符合正态，但其 μ 和 σ 均未知，若有 $S_L(\text{下限}) \leq x \leq S_U(\text{上限})$ ，产品就合格（否则不合格）。双侧抽验目的在于确保不合格品率（ p ）不大于给定的 p 值——以高概率控制的正态分布两个尾区面积分别不超过给定的 p_1 与 p_2 。

设 p_1 为 $x \leq S_L(\text{下限})$ 的不合格率允许值， p_2 为 $x \leq S_U(\text{上限})$ 的不合格率允许值，与单侧规格方案相仿，可以建立如下的判断规则：

对于任意指定的样组 n ，若希望以 90%概率保证在抽验判断为合格批中，左尾区不合格品率不超过 p_1 ，同时右尾区的 p 不超过 p_2 。查表可确定适用于此要求的LC（ $p_1=0.01$ ， $n=50$ ，有 $k_1=2.74$ ），以及（ $p_2=0.05$ ， $n=50$ ，有 $k_2=1.97$ ）。

本标准假定，测量系统的校准误差要小于所有变异源的合成，后者包括：样品制备、不同人员操作、不同仪器或设备（不同厂商、型号、硬件、电子、分析柱、抓样速率、化学处理速率、积分时间、软件计算、内部信号处理和门槛、样品有效容积、污染水平）。这种校准误差的减少可以采用苛刻的程序（不同浓度、重复测试、严格的校准可接受准则）以及通过校准验证。

本标准依据的是 ASTM D19 制订的实验室间统计程序，而 D19 正在利用室间的同一原理，即将出台室内的相应评定程序，这样做迎合了使用者的需求。故本标准在范围的界定中，提到了“也适用于室内的检出限和定量限的评定”。

本标准属于综合性基础标准，建议作为推荐性国家标准发布实施。

四、国内外现状及本标准研究特点

国内和国外都有类似的评定方法，但有的方法与本标准之间存在很多不同的技术问题，而这些技术问题都是关键性的处理方法，集中体现在以下几个代表性的标准中。

（1）国内的ICP（电感耦合等离子光谱仪）测量系统方法标准

着重推荐了 S/N （信噪比）或空白样的 3σ 或 10σ 的计算方法，这种方法非常简单，但其科学性和严谨性值得商榷。

本标准坚持认为，从 0 水平到 LD 或 LQ 水平， SD 呈显著变化，必须对偏倚进行修正。而目前的传统做法是利用 3σ 和 10σ 的单点估计，并认为 0 水平的变异等同于高水平下的变异。该单点估计方法尽管简单易理解，但容易误导，错误评定系统中的噪声和偏倚特性，会带来重大的损失代价，见图 1。

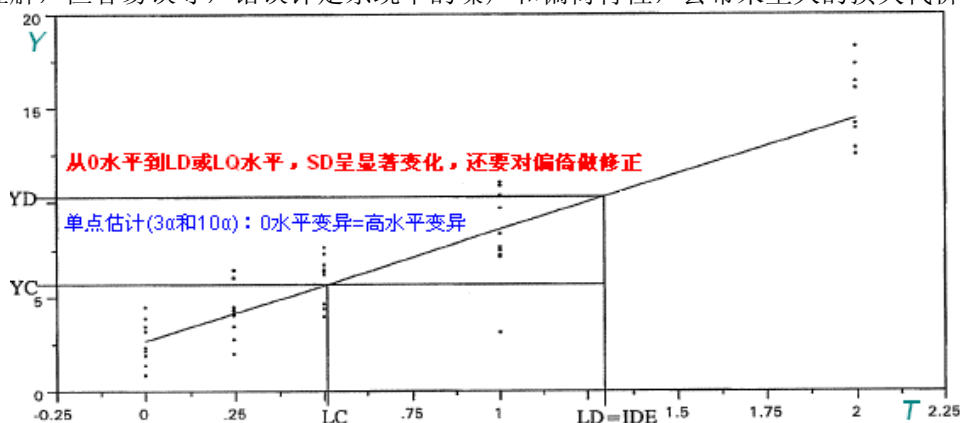


图 1 LC 和 LD 示意图

（2）美国ACS（美国化学协会）的方法程序

认为“检出限”需要 WLS （加权最小二乘）拟合，而“定量限”需要 RSD 来计算。给出了 4 种选择： S/N 的 3σ 或 10σ 、 OLS （普通最小二乘）、 WLS 拟合、以及 Eurachem 法。

鉴于理论上无法证明这 4 种方法的一致性，故 ACS 做了个实验用来进行比较。比较的结果是：

由于拟合显著（ F 比过大），常数模型则需要进行 WLS 拟合，而且认为 OLS 和 WLS 给出的结果之差很大。建议不要使用 OLS 方法， WLS 方法比较理想。尽管 S/N 方法比较简单，但碍于操作者的手法和习惯。

（3）ISO 11843 的方法程序

该系列共有以下 5 个标准

ISO 11843-1：术语和定义；

ISO 11843-2：线性校准的计算；

ISO 11843-3：非校准数据使用下响应变量临界值的测定；

ISO 11843-4：最小可检出值与给定值的比较；

ISO 11843-5：线性与非线性校准的计算。

首先，这个系列标准给出了苦涩的基本术语（状态变量、基本状态、参照状态、净状态变量、响应变量和净状态变量的临界值、净状态变量的最小可检出值，等等）。

其次，这个系列标准强调的是 $\alpha=\beta=0.05$ 的置信区间（非容忍区间），也就是说仅强调了 α 的控制，并不注重于 β 的控制。

第三，这个系列标准的假定建立在 ISO 11095 基础上（正态性和独立性），它不做 p 检验，却仅根据目测作图来判断，而且对回收模型不做系统效应检验。

最后，这个系列标准对检出限的迭代计算复杂（其中涉及到了中间辅助变量的繁琐计算）。

（4）ASTM D19 的方法程序

首先，认为属于静态评定的 3σ 和 10σ 不科学，测量系统应是依赖于水平的异方差现象，只能采用 WLS 拟合。

其次，需要同时控制 α 和 β 两个错误概率。

最后，是一种非置信的容忍区间评定（两个区间的评定结果不一样），这样 LD 起到了一个符合性

的监控作用。 LD 还使用了容差域，由此，后续的测量就会合理地纳入其中，详见图 2。

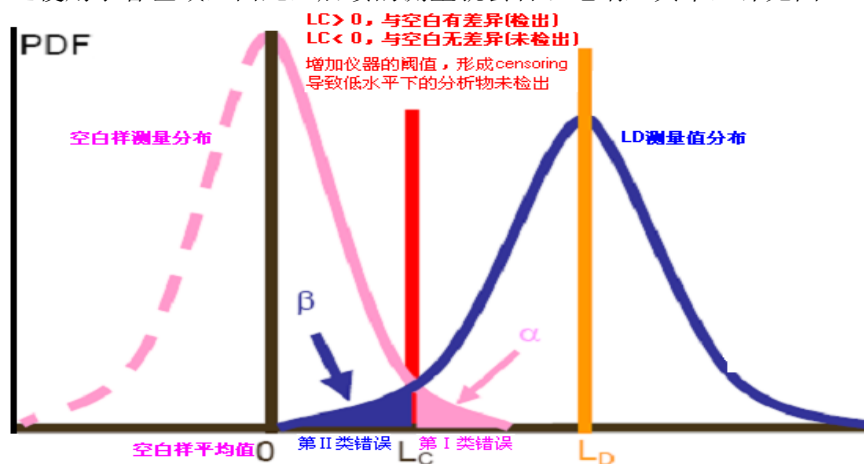


图 2 LC 和 LD 的作用

至于对定量限的理解有精密度和偏倚规定的最低校准点。 LQ 与 LD 不同，当浓度等于 LD 时，虽然检测出其分析信号证明其存在，但定量误差却很大，此时的测量结果仅具有定性分析的价值。定量限的 RSD 固定在 10%，这种 RSD 的容差固定在每个变量的低和高范围内，其估计范围是根据 R 指数曲线的转折点，见图 3。

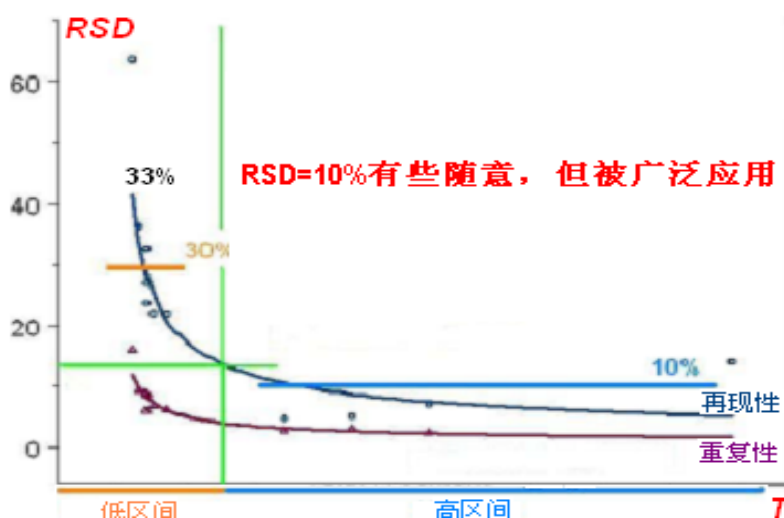


图 3 LQ 的划分点

(5) 本标准与其它方法的不同之处（括号内为其它方法）

- ① 概念定义：遵循 Currie 的原则，并规范了术语定义（比较乱）；
- ② 标准差：建立标准差随水平浓度显著变化的统计模型（认为标准差是不变化的单点校准）；
- ③ 模型检查：标准差统计模型的合理性与回收模型的假设检验（不做统计检验）；
- ④ 区间界定：给出基于 n 的 90% 置信下单侧容忍区间（仅给出置信区间）；
- ⑤ 两类错误：同时考虑到 α 和 β 的水平控制（仅给出 α 水平）；
- ⑥ 删失测量：特别强调空白水平下的斟酌测量（未做要求）；
- ⑦ 检出限和定量限：基于所选的标准差模型（直接估计为 3σ 和 10σ ）；
- ⑧ 临界值：仪器内门槛阈值的最优化（未做要求）；
- ⑨ 拟合：必须实施 WLS 拟合（只做 OLS 拟合）。

五、本标准的工作简介

- (1) 2007 年 3 月 15 日（北京）——TC261 主持召开标准制订工作会议。
- (2) 2007 年 4 月 4 日～6 日（大连）——辽宁出入境检验检疫局组织 GB/T 标准制订研讨会议，规定了参与单位和人员、以及编写宣贯材料任务。
- (3) 2007 年 6 月 4 日（北京）：TC261 下发国认标委秘函（成立 GB/T 起草组的通知）。

(4) 2007 年 10 月 10 日~14 日 (大连): 辽宁出入境检验检疫局召开 GB/T 标准草案研讨会, 对标准制订的技术内容进行了充分讨论。

(5) 2010 年 4 月 8 日~9 日 (北京) ——CNAS 邀请国内专家, 组织 GB/T 标准制订研讨会议。

(6) 2010 年 7 月 19 日~20 日 (北京) ——CNAS 邀请国内专家, 组织 GB/T 标准研讨会议。

(7) 2010 年 8 月~9 月 (大连) ——根据提出的反馈信息, 在工作组内进行详细讨论, 随后集中意见对形成的草案再次做针对性调整。

上述形成的各阶段标准文稿, 均通过电子邮件、电话、专家会议等方式来完成。专家们对标准逐条进行了反复的讨论, 提出了许多修改建议, 并在征求反馈意见的基础上, 对最初草案稿做了仔细地调整和相应的修改, 最终完成了本标准的征求意见稿。

六、标准制订的技术内容说明

(1) 关于 D6091 和 D6512 拟合在一起的设想

检出限和定量限的研究应不可分割, 尽管两者之间的数据分布及其表达方式有一些差异, 但都是用一个统计原理和分析程序。ASTM D19 正在制定的室内程序中也建议, 可以使用定量限的试验设计, 通过一组数据集来同时评定两限。

鉴于此, 在征求了相当多数的现场分析人员意见后, 认为两限的合并研究非常有助于本标准的实施。

(2) 关于术语中“统计容忍区间”的理解

这应该是一个核心的问题之一, 也是本标准区别于其它任何标准的关键所在。在这里简单解释一下置信区间 (confidence interval) 和包含区间 (tolerance interval) 两者间的区别:

置信区间有

① 如果有 T_1 和 T_2 两个统计量 ($T_1 \leq T_2$), 且 θ 为总体分布拟估计的未知量, 那么 $[T_1, T_2]$ 以一定的 $P(C)$ 包含 θ , 这样, 该区间视为 θ 的双侧置信区间。

② 在该区间内, 若 T_2 为 ∞ 或未知量的上限, 或者 T_1 为 $-\infty$ 或未知量的下限, 则该区间称为单侧置信区间或置信限 $\alpha(C)$ 。

③ 如果有 $P(T_1 \leq \theta \leq T_2) \geq 1 - \alpha$, 则该区间的置信水平 (置信系数、置信度) 为 $1 - \alpha$ 。

容忍区间有

如果有 $T_1 \leq T_2$, 那么此区间的总体分布 $F(x)$ 的 (β, γ) 称为容忍区间或容忍限 $\alpha(T)$, 其以不低于 γ 的包含概率 $P(T)$ 至少容许总体的一个确定比例 β , 即 $P[F(T_2) - F(T_1) \geq \beta] \geq \gamma$, 其中的 T_2 和 T_1 分别称为上、下限。

更进一步的分析有

对于一定的 $\alpha(C)$ 来说, 随着 n 的增加, $P(C)$ 也增大, 当 $n \rightarrow \infty$, $P \rightarrow 1$, 或 $\alpha(C) \rightarrow 0$ 。

在涉及到单次测量结果的不确定度时, $P(C)$ 和 $P(T)$ 混用, 若按大数法则以概率而收敛于某个 $P(C)$, $P(C)$ 可理解为 $|x - \mu|$ 不超过 $k(C)s = k(C)\sigma / \sqrt{n}$ 的概率。于是, 在任一次的单测量中, 有 $P\{|x - \mu| \leq \alpha(T)\} \geq P(T)$ 。

针对可测的不确定度, 当通过一系列等精度测量来确定误差的分布律时, 可以考虑确 $P(C)$ 。但对于不可测的不确定度, 这种 P 往往是不清楚的, 其估计常常带有相当浓厚的主观臆断的色彩。

(3) 关于“LC和LD”的理解

实际上, 这是一个弃真和存伪概率的释义。

空白样测量的 LC , 可视为“未检出概率” (负概率), 即 $1 - \alpha$ 是 99%, 其对应的“错误正概率” (错误检出概率) α 是 1%。 LD 的“检出概率” $1 - \beta$ 是 95%, 同时空白样测量时的 YD 有 1% 超出 YC 。

在 LC 点, 如果数据在 α 控制下过多删失, 则报告 β 的控制会过大, 所以 LD 必须在较高水平下达到一个合理的 β 来进行估计。传统的方法仅强调 α 的控制, 尽管某些情况下比较适宜, 但有时在许多重要场合下, β 的控制也同样重要。本标准将 β 减少到可接受水平, 强调了 β 的控制风险, 有助于 α 错误率的减少。

(4) 关于“方案设计”的考虑

校准母液的制备来自于 RM 样品或 CRM 稀释溶液, 配备的梯度溶液必须当天用完以获取最佳结果。如果以前的质控活动未发现问题, 则母液可半年配制一次。

最常用的方法是使用浓度已知的标准样品，鉴于这种 *RM* 不易获得，则可以自己配备添加样品（视为浓度已知样品），对其组分量进行测定。

建议水平下给出至少 6 次独立重复测量。

如果 0 浓度没有包含在内，应尽可能选择一个靠近 0 的浓度。

在此痕量分析的范围内，要求水平的分布一定要均匀。尤其在检出限和定量限的研究中，仔细考虑实验室对删失的测量。

避免选择超出痕量分析范围的过高水平数据，因为在高浓度下的响应会变平，导致不同的模型，这样会影响 *LC* 和 *LD* 的测定。

许多情况下，仪器的调 0 都是人为控制。作为日常的痕量分析，目的是希望得到实际中大量的未删失、无偏的、低水平数据，所以在痕量测试中，任何内部测量系统的门槛限值，都应该予以减少和最佳化。

（5）关于“干扰和数据”的筛选

要求所有的试剂、溶剂、载气的纯度高，以减少干扰带来的问题。

污染物会造成基体的干扰，而基体干扰的程度可能随样品源而显著不同，这些样品源完全来自于样品基体的变异。

方法的干扰可能来自于溶剂、试剂和容器中的污染，也可能其它装置产生不同人员的调整或基线过高所致。应该在样品测试的同一条件下对试剂空白进行分析，以表明排除干扰。

实验室提交的结果不能有背景扣减，应报告所用每种基质的背景水平。

方法空白若大于 *LC*，就需要停止样品的分析，直至排除干扰因素。

通常情况下对于某单个水平下，若有 1/3 以上报出的数据属于非数值，则应该将该水平自方法研究中剔除，所以应该特别对实验室的删失做法予以注意。

只要实际测试给出 0 值或负值，就应该予以报告，不能进行删失，因为以“小于”或“大于”的形式来报告结果，会导致统计计算失败。也不能用 0 值来替代“小于”或其它无法定量的测试结果。

（6）关于“标准差模型”的拟合和选择

数学和作图关系对于实验室间比对结果的研究非常重要，必须给出精密度和偏倚随浓度水平变化的通用模型关系式，即在所研究的范围内来评估特定浓度下的精密度或偏倚。除非在测试方法中已经给出了这种关系式，否则就必须来找到这种数学关系式，以便对实验室间比对给出的统计量进行最佳拟合。

原文给出了三种选择模型（检出限有：常数、直线和指数；定量限有：常数、直线和混合），这些模型的拟合和选择与化学知识、方法掌握、系统熟悉等有很大关系，但是本标准更注重的是，利用统计方法来对这些模型进行拟合和选择判断。

本标准坚持认为，标准差随浓度变化是来自于方法原理和测量系统，所以，在痕量分析中，常数假设是一种错误的假设，搞不好会影响到 *LC* 和 *LD* 的估计。由于直线模型具有广泛的使用性，则本标准极力推荐直线模型的使用。

鉴于模型选择的显著性统计检验涉及到过多的方差分析描述，故本标准在相应环节中引用了有关统计标准以供参考，例如，表 1 为 D6091 的数据集统计量，图 4 为表 1 数据集的 *p* 值检验。

同理，指数模型和混合模型的拟合比较复杂，故安排在本标准的资料性附录中。

表 1 系列水平数据集的标准差统计汇总

T_k	不同实验室(不同人员)测量结果 y_{nk} , ppb										平均值	标准差	预测值
ppb	1	2	3	4	5	6	7	8	9	10	$\bar{y}_k$	s_k	拟合 $\hat{s}_k$
0.0	1.41	3.94	2.22	3.48	1.96	0.92	2.17	2.36	4.50	3.26	2.622	1.138	1.089
0.25	4.10	3.51	4.07	4.34	4.54	2.76	2.03	4.13	6.06	6.47	4.201	1.335	1.328
0.50	3.97	7.34	6.41	6.25	6.38	7.64	4.67	6.74	4.38	6.48	6.026	1.254	1.567
1.0	7.54	7.68	8.38	7.14	3.12	10.97	11.15	10.44	9.73	7.27	8.342	2.405	2.046
2.0	8.20	13.97	12.88	18.31	16.47	16.06	12.56	14.21	13.96	17.37	14.40	2.900	3.003
0	s_k vs. T_k 进行回归: 截距=1.08855 斜率=0.95701												

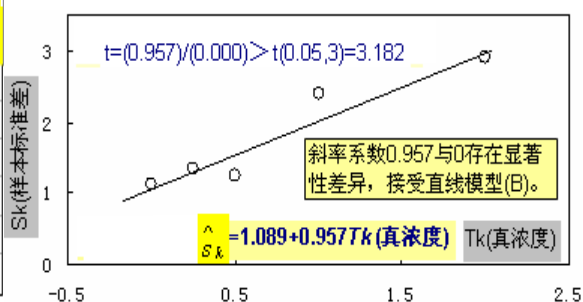


图 4 SD 对 *T* 的 OLS 拟合与直线斜率的 *t* 零假设

(7) 模型R的拟合与检查

模型R即为平均回收模型，用于LC、LD和LQ的估计。模型R是一种偏倚修正函数式，本标准要求使用WLS来加权拟合，其目的给予低水平浓度更多的权重影响，能更好地评价低水平下的偏倚。

本标准提倡使用的直线模型的回归，避免OLS的拟合，原因在于OLS假定，在独立变量的整个区间内R是常数变异。所使用的WLS回收拟合，可以防止高浓度对直线拟合的影响效应过大。

通常认为，截距大于0的正偏和的斜率不等于1表示为回收不完美，然而本标准认为，截距0和斜率1并非正确的假设，而要求进行模型拟合的失拟误差检验（详见GB/T 22554）。

例如，表2采用了美国PROXI多道光谱仪ICP-AES法，给出汽油中钒元素的校准曲线，以评价该平均回收模型拟合是否正确。表中为测量系统所测的谱线强度（ug/g），以及函数式转换后的浓度重复测量值，基于OLS拟合的截距和斜率非常完美。

表2 汽油中钒测定的校准曲线拟合

浓度	谱线强度(I)，重复次数			谱线平均值	C_n
C_n	I_1	I_2	I_3	$\bar{I}_n$	vs. $\bar{I}_n$
0	164507	166202	166240	165650	截距
1	1756046	1767907	1752749	1758901	112940
5	7570117	7699475	7717090	7662227	斜率
10	15681042	15839213	16021304	15847186	1561388

被测浓度(C_{nk})，重复次数			平均值	标准差	DEVSQ计算	C_n
C_{n1}	C_{n2}	C_{n3}	$\bar{C}_n$	SD	$E(C_{nk}-\bar{C}_n)^2$	vs. $\bar{C}_n$
0.0330	0.0341	0.0341	0.0338	0.0006339	0.00000080	$\beta_0 =$
1.0523	1.0599	1.0502	1.0542	0.005106	0.00005214	0.00000
4.7760	4.8588	4.8701	4.8350	0.051400	0.00528383	$\beta_1 =$
9.9707	10.0720	10.1886	10.0771	0.109051	0.02378429	1.00000

可是经过残差值的拟合后发现，其中的失拟误差要大于试验误差（见表3），预示着OLS拟合的截距和斜率并不完美，说明线性函数的模型拟合错误。

表3 校准曲线拟合的残差计算与失拟误差检验

0	C_{nk} 拟合	残差值 $e_{nk}=C_{nk}-C_{nk}$ 拟合		
C_n	基于校准式	e_{n1}	e_{n2}	e_{n3}
0	0.00000	0.03303	0.03411	0.03414
1	1.00000	0.05234	0.05993	0.05023
5	5.00000	-0.22401	-0.14116	-0.12988
10	10.00000	-0.02932	0.07198	0.18861

$SST = \sum \sum (C_{nk} - \bar{C})^2 = 186.141$	$SSP = \sum \sum (C_{nk} - \bar{C}_n)^2 = 0.0291$
$SSL = SSE - SSP = 0.1117$	$\sigma_L^2 = SSL / (n-2) = 0.05587$
$SSR = SST - SSE = 186.00$	$\sigma_p^2 = SSP / [n(k-1)] = 0.00364$
$SSE = \sum \sum (e_{nk})^2 = 0.1409$	$(\sigma_L^2) / (\sigma_p^2) = 15.348 > F(0.95, 2, 8) = 4.46$
残差 σ^2 估值 $(nk-2) = 0.0141$	残差 σ 估值 = 0.1187

然而，有人会认为这些数据集可能存在异方差现象，需要WLS来进行拟合，经过检验后确实需要数据变换。但是变换后的残差拟合，依然发现模型中的失拟项显著（见表4），再次证明了测量系统数据集模型拟合的错误，可见，回归系数斜率不能作为线性拟合好坏的评判依据。

表4 校准曲线WLS拟合后的残差计算与失拟误差检验

0.0111	-0.002830	权值 w_n	z_{n1}	z_{n2}	z_{n3}	$z_{n\bullet}$	$\sum (z_{nk} - z_{n\bullet})^2$
0.0004	0.0024895	0	0	0	0	0	0
0.997	0.0034927	1.00	1.05234	1.05993	1.05023	1.0542	0.0000521
625.6	2	0.20	0.95520	0.97177	0.97402	0.9670	0.0002114
0.008	0.0000244	0.10	0.99707	1.00720	1.01886	1.0077	0.0002378
$t=25 > t_{0.01,2}=9.9$		$y_0 = 0.9777$	$y_1 = 0.0736$				
结论：线性拟合斜率与0存在显著差异（出现异方差），需要数据变换。		$WSST = \sum \sum (z_{nk} - \bar{z})^2 = 0.0119$					
		$WSSP = \sum \sum (z_{nk} - z_{n\bullet})^2 = 0.0005$					
		$\sigma_p^2 = WSSP / [n(k-1)] = 0.0001$					

y_n 拟合	z_n 拟合	u_{n1}	u_{n2}	u_{n3}
0.07361236	0	0	0	0
1.051337516	1.05134	0.001000	0.008596	-0.00111
4.96223814	0.99245	-0.037249	-0.02068	-0.01842
9.850863919	0.98509	0.011982	0.022112	0.033774
$WSSL = WSSE - WSSP =$		$WSSE = \sum \sum (u_{nk})^2 = 0.004004$		
0.003503		加权残差 τ^2 估值 = 0.00040		
$\sigma_L^2 = WSSL / (n-2) =$		加权残差 τ 估值 = 0.02001		
0.001751		$(\sigma_L^2) / (\sigma_p^2) = 27.95$		

(8) 检出限和定量限的计算与报告

关于系数k的选择是基于数据质量目标，有 $\alpha=0.01$ 和 $\alpha=0.05$ 的样本容差域区间90%计算。求得YC计算值后，再通过模型R给出LC。如果使用者对YD值感兴趣，也可给出相应的YD值。

初始值的精确起点不是一个关键的问题，而是看它如何来收敛，即两个起始点之间的差异是否最小。所有的数据都可以用在基线风险的评估中，只要对这些数据的不确定度进行清楚描述。

在所规定的校准线性范围内的方法性能最低限LQ，在界定的容忍区间内，可以获得超出LQ限的定量结果，通常视为2~5倍于LD，处于两者之间可视为“定量未被检出”。

方法检出限包括样品的制备效应，相比较而言，所估计的程序要比仪器检出限准确得多。

风险评估者应清楚所需的浓度或初步的补偿目标。当分析物的浓度大于LD时，此时在容忍区间内可使用“未检出”和“检出”的结果。如果两者叠盖，则会有可能产生 α 和 β 错误率。当浓度水平小于LD时，有证据预示存在可能的 β 错误率，仅表明“检出”结果的有效性。

分析物浓度超出LQ，适用于样品的稀释分析，但必须在线性限内，否则仪器会由于饱和而导致仪器响应下降，由此给出的浓度报告值会低于样品中存在的实际量。

图5给出了表1（D6091数据集）的WLS拟合模型选择。

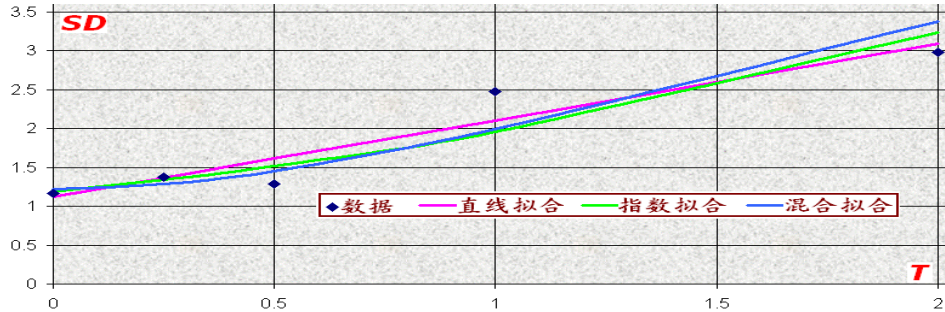


图5 不同实验室（参加人员）的D6091数据集WLS拟合下模型选择比较

图5大致给出直线模型拟合的一个合理性，详细的统计量见表5。

表5 不同实验室（参加人员）的D6091数据集模型选择下的LD/LQ评定比较

	p	平方和	WLS 拟合 a/b	WLS模型 g/h	LC	LD	$IQE_{10\%}$	$IQE_{20\%}$	$IQE_{30\%}$
常数	0.01281 <0.05	2.673	2.765 5.804	1.119 0.984	0.959	1.650	3.438	1.719	1.146
直线	0.01281 <0.05	0.2540 (min)	2.724 5.872	1.119 0.984	0.532	1.429	—	5.865	1.438
指数	0.01951 <0.05	0.3833	2.738 5.859	1.184 0.501	0.564	4.118	—	—	1.283
混合	0.7064 >0.05	0.4219 (max)	2.747 5.863	1.221 1.570	0.581	1.482	—	—	1.541

图6给出了表6（D6512数据集）的WLS拟合模型选择。

表6 不同实验室（参加人员）的D6512数据集统计

T_k	不同实验室（不同人员）测量结果 y_{nk} , ppb										标准差	标准差	自然对数	混合模型	WLS	$(T_k)^2$	$(T_k)^2$	拟合- T_k^2	Qq_k
ppb	1	2	3	4	5	6	7	8	9	10	s_k	s_k 调整	$\ln s_k$ 调整	s_k 拟合	w_k	$(T_k)^2$	拟合预测	q_k	
0	-0.105	0.263	0.293	0.187	0.106	0.329	0.080	0.524	0.278	0.206	0.168	0.1728	-1.756	0.2334	18.363	0.0	-13.029	-13.029	0.168
0.5	0.354	0.724	0.682	0.327	0.527	0.868	0.730	0.434	0.794	0.642	0.188	0.1931	-1.645	0.2247	19.811	0.25	-7.2	-7.453	0.096
1	1.241	0.668	1.200	1.370	1.106	0.964	0.949	1.421	1.032	1.134	0.221	0.2270	-1.483	0.2224	20.21	1.0	-1.4	-2.376	0.031
2	2.174	2.388	2.153	2.366	2.306	2.309	1.663	2.841	1.933	1.809	0.335	0.3447	-1.065	0.2374	17.747	4.0	10.3	6.277	-0.08
4	3.660	3.734	3.167	3.578	4.278	3.383	3.873	4.479	3.919	3.856	0.389	0.3995	-0.917	0.3448	8.4113	16.0	33.6	17.582	-0.23
8	6.592	7.520	6.822	7.751	7.771	7.296	8.578	6.863	7.840	8.821	0.732	0.7522	-0.285	0.8689	1.3216	64.0	80.2	16.194	-0.21
12	9.496	9.081	13.942	10.547	9.324	13.148	10.994	11.774	12.320	13.521	1.801	1.8518	0.6162	1.8086	0.3057	144	126.8	-17.194	0.222

图6大致给出表6数据集指数模型拟合的一个合理性，详细的统计量见表7。

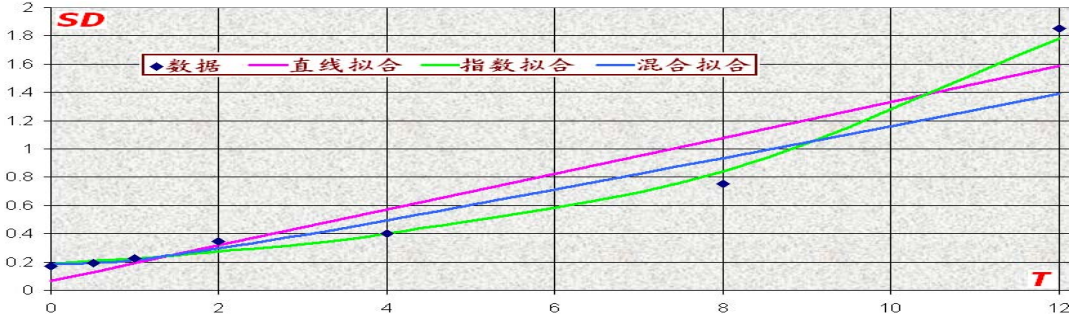


图6 不同实验室（参加人员）的D6512数据集WLS拟合下模型选择比较

表 7 不同实验室（参加人员）的 D6512 数据集模型选择下的 $LDILQ$ 评定比较

	p	平方和	WLS 拟合 a/b	WLSD 模型 g/h	LC	LD	$IQE_{10\%}$	$IQE_{20\%}$	$IQE_{30\%}$
常数	0.01217 <0.05	2.172 (max)	0.1874 0.9312	——	2.288	3.929	8.504	4.252	2.835
直线	0.01217 <0.05	0.2249	0.2042 0.9928	0.065 0.127	0.189	0.442	——	1.123	0.433
指数	0.000016 <0.05	0.0186 (min)	0.1998 0.9265	0.1885 0.1871	0.547	0.982	4.961	1.296	0.785
混合	0.009557 <0.05	0.2605	0.1940 0.9306	0.1841 0.1146	0.532	0.979	——	1.255	0.723