

实验室间比对能力验证中的两种稳健统计技术探讨

邢小茹, 马小爽, 田文, 吴忠祥 (环境保护部标准样品研究所, 北京 100029)

摘要: 对环境监测实验室能力验证中两种稳健数据统计技术以及 z 比分数评价准则进行了探讨, 并介绍了如何使用 Excel 电子表格软件实现数据的统计处理。四分位数法和迭代法均可作为实验室比对能力验证的数据处理技术, 相对于四分位数稳健统计方法, 迭代法是一种更加温和且可靠的稳健统计方法, 但不适用于分割水平设计的能力验证计划。

关键词: 稳健统计方法; 四分位数法; 迭代法; z 比分数

中图分类号: X830.5 **文献标识码:** A **文章编号:** 1002-6002(2011)04-0004-05

Two Robust Statistic Techniques in Proficiency Testing by Interlaboratory Comparisons

XING Xiao-ru, et al. (Reference Materials Institute of SEPA, Beijing 100029, China)

Abstract: Two robust statistic techniques and z-score in proficiency testing of environmental monitoring laboratories were discussed and Excel implementation. Quartile and iteration robust statistic techniques both can be used in data analysis of proficiency testing by interlaboratory comparisons. Compare to quartile robust statistic techniques, iteration technique is more mild and reliable, but it is not suitable for proficiency testing with split-level test sample.

Key words: Robust statistic technique; Quartile; Iteration; Z-score

实验室间比对能力验证是通过实验室间测量比对评价实验室进行某种特定检测或测量的能力, 以及对实验室持续能力的监控。环境监测实验室的能力验证工作对保证环境监测数据的准确性和可靠性, 确保环境质量管理工作的顺利进行非常重要。中国合格评定国家认可委员会 (CNAS) 要求已认可的实验室和申请认可的实验室必须参加能力验证活动^[1]。实验室通过参加能力验证计划可客观评价其数据的准确性和可靠性, 也可从中找出问题并采取相关的补救措施, 对实验室的质量控制和保证起到补充和完善的作用^[2]。近年来, 由管理部门 (如 CNAS) 以及各行业组织开展的能力验证计划越来越多, 如何才能达到能力验证的目的并对实验室的检测能力进行客观评价, 除要求参加实验室提供准确可靠的实验结果外, 实验数据的统计分析技术以及参考值的评估方法也非常重要。

传统的统计方法对数据正态分布的假定具有很强的依赖性, 而实际数据大多不能服从正态分布, 最多是近似的正态分布。因此, 在处理实际问题时, 要对传统的统计量进行稳健化。稳健统计的发展已有 50 多年的历史, 但对实验室能力验证

分析来说是一种新的统计技术, 目前实验室能力验证计划也越来越多地采用这种数据处理方式。所谓稳健统计技术就是统计方法能够很好而且合理处理假定模型, 当模型有较少偏离时, 其结果也应该只遭到少许破坏, 当模型有较大偏离时, 结果也不应该遭到破坏性的影响^[3]。稳健统计实际上就是减少正态分布假设或异常值 (离群值) 对总体的影响。能力验证中的稳健统计方法主要有 ILAC 的四分位数稳健统计以及 ISO (ISO 13528) 的迭代稳健统计。目前, 国内的能力验证计划主要采用四分位数稳健统计技术, 而 APLAC 要求能力验证计划采用 ISO 13528 的迭代稳健统计技术。ISO 13528 的国标化工作已经开始, 鉴于此, 本文对这两种稳健统计技术的原理及应用进行详细讨论, 希望能对国内的能力验证结果评价工作有所帮助。

1 两种稳健统计方法的基本理论

1.1 四分位数稳健统计

四分位数稳健统计使用中位值和标准四分位间距分别代替传统统计方法中的平均值和标准偏

收稿日期: 2010-12-08; 修订日期: 2011-02-23

基金项目: 中国合格评定国家认可委员会立项并提供资助 (CNASTO497)

作者简介: 邢小茹 (1975 -), 女, 陕西富平人, 博士。

差作为总体的估计,即对检测结果总体参数的估计。四分位数稳健统计方法的统计参数包括:结果数量(总统计量 N)、中位值(M)、标准四分位间距($NIQR$)、稳健变异系数(CV)、最小值(MIN)、最大值(MAX)以及变动范围(R)。四分位间距为上四分位值(75%分位值 Q_3)与下四分位值(25%分位值 Q_1)的差,标准四分位间距 $NIQR = 0.7413 \times (Q_3 - Q_1)$,相当于一个标准差。稳健变异系数 $CV = NIQR/M \times 100\%$,它表示测试结果之间的变异性。变动范围为最大值与最小值之差,即 $R = MAX - MIN$ 。

1.2 迭代稳健统计

ISO 13528(实验室间比对能力验证中的统计方法)中涉及的迭代稳健统计分析包括 A 算法和 S 算法,这两种算法来源于 ISO 5725-5(测量方法和测量结果的精确性——第 5 部分标准测量方法精度测定的迭代法)^[4]。其中 A 算法用于估计数据总体的稳健平均值和标准偏差;S 算法用于合并实验室内部给出的标准偏差或变动范围。如果实验室均给出了检测数据的实验室内标准偏差或者数据的变动范围,则可以用 S 算法给出数据再现性的标准偏差。能力验证工作中确定检测变量稳健平均值和标准偏差的迭代法大多为 A 算法,具体步骤如下:

假设有 p 个实验室数据,首先按照递增的顺序排序: $x_1, x_2, \dots, x_i, \dots, x_p$

设这组数据的稳健平均值和稳健标准偏差分别为 x^* 和 s^* 。

x^* 和 s^* 的初始值如下:

$$x^* = x_i \text{ 的中位值 } (i = 1, 2, 3, \dots, p) \quad (1)$$

$$s^* = 1.483 \times (|x_i - x^*| \text{ 的中位值})$$

$$(i = 1, 2, 3, \dots, p) \quad (2)$$

x^* 和 s^* 的迭代更新值如下,先计算:

$$\delta = 1.5 \times s^* \quad (3)$$

针对每个 $x_i (i = 1, 2, 3, \dots, p)$, 计算:

$$x_i^* = \begin{cases} x^* - \delta, & \text{如果 } x_i < x^* - \delta \\ x^* + \delta, & \text{如果 } x_i > x^* + \delta \\ \text{其它} & \end{cases} \quad (4)$$

用下列公式计算新的 x^* 和 s^* :

$$x^* = \sum x_i^* / p \quad (5)$$

$$s^* = 1.134 \sqrt{\sum (x_i^* - x^*)^2 / (p - 1)} \quad (6)$$

至此完成了第一次迭代。第二次迭代从公

式(3)开始至公式(6)结束,依次类推,直至新的稳健平均值 x^* 和稳健的标准偏差 s^* 收敛至小数点后第三位数保持不变,最后得到的 x^* 和 s^* 即为该组数据的稳健平均值和标准偏差^[4]。

2 应用举例

2009 年 11 月至 2010 年 2 月, CNAS 与国家环境保护部标准样品研究所合作开展了“水中铅、镉和砷检测能力验证计划”(CNAS T0497)。该项目共有 460 家各行各业的实验室参加,其中环保系统的实验室有 121 家,主要为各省市环境监测站。本文以此项目 Pb 的部分测试数据为例,具体介绍如何用 Excel 软件实现以上两种稳健统计方法的计算。

2.1 四分位数稳健统计的 Excel 计算

表 1 为 CNAS T0497 能力验证项目中水中 Pb 的部分实验室检测数据的处理过程,表中的实验室代码只适用于本文,与项目实施过程具体的实验室代码无关。四分位数稳健统计方法所涉及的统计参数及计算公式均在表中列出。四分位法稳健统计得出 Pb 浓度为 1.095mg/L,标准偏差为 0.039mg/L。

表 1 CNAS T0497 能力验证项目中部分 Pb 检测结果的四分位数稳健统计分析

	A	B	C	D
1	实验室代码	Pb 浓度		
5	01	1.08	N	24
6	02	1.07	M	1.095
7	03	1.02	Q_3	1.123
8	04	1.13	Q_1	1.070
9	05	1.20	$NIQR$	0.039
10	06	1.07	MIN	0.930
11	07	1.09	MAX	1.20
12	08	1.15	稳健 CV	3.60%
13	09	1.02	R	0.270
14	10	1.10		
15	11	1.13		
16	12	1.10		
17	13	1.20		
18	14	1.12		
19	15	1.16		
20	16	0.930		
21	17	1.08		
22	18	1.10		
23	19	1.11		
24	20	0.988		
25	21	1.09		
26	22	1.07		
27	23	1.02		
28	24	1.10		

表 1 中涉及的函数及公式如下:

结果数量 N :	$D5 = \text{Count}(B5:B28)$	稳健变异系数 CV :	$D12 = D9/D6 * 100\%$
中位值 M :	$D6 = \text{Median}(B5:B28)$	变动范围 R :	$D13 = D11 - D10$
上四分位数 Q_3 :	$D7 = \text{Quartile}(B5:B28,3)$	2.2 迭代稳健统计的 Excel 计算	
下四分位数 Q_1 :	$D8 = \text{Quartile}(B5:B28,1)$	表2列出了迭代稳健统计方法的具体计算步骤,为了便于比较,使用的原始数据与表1相同,也是CNAS T0497能力验证计划中Pb的部分检测结果。	
标准四分位间距 $NIQR$:	$D9 = 0.7413 * (D7 - D8)$		
最小值 MIN :	$D10 = \text{Min}(B5:B28)$		
最大值 MAX :	$D11 = \text{MAX}(B5:B28)$		

表2 CNAS T0497能力验证项目中部分Pb检测结果的迭代稳健统计分析

	A	B	C	D	E	F	G	H	I
1	代码	0次	$ x - x^* $	1次	2次	3次	4次	5次	6次
2	$1.5s^*$			0.056	0.065	0.073	0.080	0.083	0.085
3	$x^* - 1.5s^*$			1.039	1.028	1.019	1.011	1.007	1.006
4	$x^* + 1.5s^*$			1.151	1.158	1.165	1.171	1.174	1.175
5	01	1.08	0.015	1.080	1.080	1.080	1.080	1.080	1.080
6	02	1.07	0.025	1.070	1.070	1.070	1.070	1.070	1.070
7	03	1.02	0.075	1.039	1.028	1.020	1.020	1.020	1.020
8	04	1.13	0.035	1.130	1.130	1.130	1.130	1.130	1.130
9	05	1.20	0.105	1.151	1.158	1.165	1.171	1.174	1.175
10	06	1.07	0.025	1.070	1.070	1.070	1.070	1.070	1.070
11	07	1.09	0.005	1.090	1.090	1.090	1.090	1.090	1.090
12	08	1.15	0.055	1.150	1.150	1.150	1.150	1.150	1.150
13	09	1.02	0.075	1.039	1.028	1.020	1.020	1.020	1.020
14	10	1.10	0.005	1.100	1.100	1.100	1.100	1.100	1.100
15	11	1.13	0.035	1.130	1.130	1.130	1.130	1.130	1.130
16	12	1.10	0.005	1.100	1.100	1.100	1.100	1.100	1.100
17	13	1.20	0.105	1.151	1.158	1.165	1.171	1.174	1.175
18	14	1.12	0.025	1.120	1.120	1.120	1.120	1.120	1.120
19	15	1.16	0.065	1.151	1.158	1.160	1.160	1.160	1.160
20	16	0.930	0.165	1.039	1.028	1.019	1.011	1.007	1.006
21	17	1.08	0.015	1.080	1.080	1.080	1.080	1.080	1.080
22	18	1.10	0.005	1.100	1.100	1.100	1.100	1.100	1.100
23	19	1.11	0.015	1.110	1.110	1.110	1.110	1.110	1.110
24	20	0.988	0.107	1.039	1.028	1.019	1.011	1.007	1.006
25	21	1.09	0.005	1.090	1.090	1.090	1.090	1.090	1.090
26	22	1.07	0.025	1.070	1.070	1.070	1.070	1.070	1.070
27	23	1.02	0.075	1.039	1.028	1.020	1.020	1.020	1.020
28	24	1.10	0.005	1.100	1.100	1.100	1.100	1.100	1.100
29	Average	1.089		1.093	1.092	1.091	1.091	1.091	1.091
30	SD	0.062		0.038	0.043	0.047	0.049	0.050	0.050
31	New x^*	1.095		1.093	1.092	1.091	1.091	1.091	1.091
32	New s^*	0.037		0.043	0.049	0.053	0.056	0.057	0.057

可以看出,第5次迭代和第六次迭代得出的新的稳健平均值(x^*)和稳健标准偏差(s^*)小数点后第三位保持不变,因此可以结束迭代计算。迭代法得出的稳健统计结果Pb浓度为1.091 mg/L,标准偏差为0.057 mg/L。在Excel表格中具体的计算公式如表3所示,步骤如下:

第一步:将实验室代码及其检测结果输入Excel表格中(表2的A列和B列的第5到28行)。表中“0次”迭代列为参加实验室的检测结果,计算该列数据的算术平均值和标准偏差,结果

显示在单元格B29和B30中。计算B5到B28的中位值,结果显示在B31中,即为初始的 x^* 。C5~C28分别为B5~B28与B31之差的绝对值,其中位值的1.483倍为初始的 s^* ,列于B32中。

第二步:计算“1次”迭代的迭代值,即D2、D3和D4。比较B5~B28的值与D3和D4的大小,如果小于D3则用D3取代,如果大于D4则用D4取代,否则保持原数据不变。计算1次迭代的算术均值和标准偏差,列于D29~D31中。新 x^* 为D5~D28的算术均值,新 s^* 为D5~D28的标准

偏差乘以 1.134。至此完成了第一次迭代。

第三步：按照与第二步相同的步骤，依次类推，完成后面的迭代计算，直至 x^* 和 s^* 的第三位小数不变为止。最后得出的 x^* 和 s^* 即为稳健的平均值和标准偏差。

3 稳健的 z 比分数

3.1 z 比分数的计算

在实验室间比对能力验证计划中，通常需要将参加实验室的检测结果转换为一个能力统计量，其目的是通过与能力标准比较来测评偏离指定值的程度。对参加实验室来说，能力测评应该有意义。因此，测评应该与检测的需求有关联且易于理解，或者是特定领域内的传统方法。可以使用的能力统计量包括偏差、偏差百分率、百分数或秩、 z 比分数、 E_n （尤其用于方法比对计划）等。其中稳健的 z 比分数法是国际上普遍采用的能力统计量，对于单一样品的检测结果， z 比分数的计算公式如下：

$$z = \frac{x - X}{s} \quad (7)$$

式中， x 为参加实验室的检测结果；可根据参加实验室的结果导出 X 和 s ，对于四分位数稳健统计方法， X 为中位值 M ， s 为标准四分位间距 $NIQR$ ；对于迭代稳健统计方法， X 为末次迭代的 x^* ， s 为末次迭代的 s^* 。

z 比分数反映了参加实验室的检测结果与稳健值的偏离程度。 z 比分数越接近于 0，表示该实验室的检测结果偏离程度越小，反之，偏离程度越大，实验室的结果越趋于异常。 z 比分数评价参加实验室能力的准则：如果 z 值大于 3 或者小于 -3，表示参加实验室的结果为离群值，是不满意结果，实验室应查找原因，制订有效的纠正措施，并进行复查；如果 z 值大于 2 或者小于 -2，则表示实验室的结果可疑，应引起重视，查找原因；如果 z 值小于 2 或者大于 -2，则表示实验室的结果为满意结果，实验室对参加能力验证计划的项目具有良好的检测能力。

四分位数稳健统计法还可以对分割水平样品的检测结果进行分析和评价，其稳健 z 比分数又可以分为实验室间的 z 比分数 (ZB) 和实验室内的 z 比分数 (ZW)。其计算方法见表 3。

表 3 CNAS T0497 能力验证项目中部分 Pb 检测结果的迭代稳健统计 Excel 公式

A	B	C	D	E	F	...	
1	代码	0 次	$ x - x^* $	1 次	2 次	3 次	...
2	$1.5s^*$		$= 1.5 * D32$	$= 1.5 * D32$	$= 1.5 * E32$	$= 1.5 * F32$...
3	$x^* - 1.5s^*$		$= B31 - D2$	$= D31 - E2$	$= E31 - F2$	$= F31 - G2$...
4	$x^* + 1.5s^*$		$= B31 + D2$	$= D31 + E2$	$= E31 + F2$	$= F31 + G2$...
5	01	1.08	$= ABS(B5 - B$31)$	$= IF(B5 < D$3, D$3, IF(B5 > D$4, D$4, B5))$	$= IF(B5 < E$3, E$3, IF(B5 > E$4, E$4, B5))$	$= IF(B5 < F$3, F$3, IF(B5 > F$4, F$4, B5))$...
6	02	1.07	$= ABS(B6 - B$31)$	$= IF(B6 < D$3, D$3, IF(B6 > D$4, D$4, B6))$	$= IF(B6 < E$3, E$3, IF(B6 > E$4, E$4, B6))$	$= IF(B6 < F$3, F$3, IF(B6 > F$4, F$4, B6))$...
7	03	1.02	$= ABS(B7 - B$31)$	$= IF(B7 < D$3, D$3, IF(B7 > D$4, D$4, B7))$	$= IF(B7 < E$3, E$3, IF(B7 > E$4, E$4, B7))$	$= IF(B7 < F$3, F$3, IF(B7 > F$4, F$4, B7))$...
...
28	24	1.1	$= ABS(B28 - B$31)$	$= IF(B28 < D$3, D$3, IF(B28 > D$4, D$4, B28))$	$= IF(B28 < E$3, E$3, IF(B28 > E$4, E$4, B28))$	$= IF(B28 < F$3, F$3, IF(B28 > F$4, F$4, B28))$...
29	average	$= AVERAGE(B5:B28)$	$= AVERAGE(D5:D28)$	$= AVERAGE(E5:E28)$	$= AVERAGE(F5:F28)$	$= AVERAGE(G5:G28)$...
30	stdev	$= STDEV(B5:B28)$	$= STDEV(D5:D28)$	$= STDEV(E5:E28)$	$= STDEV(F5:F28)$	$= STDEV(G5:G28)$...
31	new x^*	$= MEDIAN(B5:B28)$	$= AVERAGE(D5:D28)$	$= AVERAGE(E5:E28)$	$= AVERAGE(F5:F28)$	$= AVERAGE(G5:G28)$...
32	new s^*	$= 1.483 * MEDIAN(C5:C28)$	$= 1.134 * D30$	$= 1.134 * E30$	$= 1.134 * F30$	$= 1.134 * G30$...

假设有两个样品的检测结果分别为 A 和 B , 则其标准化和 (S) 和标准化差 (D) 分别为

$$S = (A + B) / \sqrt{2}$$

$$D = |A - B| / \sqrt{2}$$

ZB 和 ZW 的计算公式分别为

$$ZB = (S - M(S)) / NIQR(S)$$

$$ZW = (S - M(D)) / NIQR(D)$$

式中, $M(S)$ 和 $NIQR(S)$ 分别为标准化和的中位值和标准四分位间距, $M(D)$ 和 $NIQR(D)$ 分别为标准化差的 中位值和标准四分位间距。

对于采用分割水平设计的能力验证计划, ZB 和 ZW 分别反映实验室的系统误差和随机误差。 $|ZB| > 3$ 说明实验室系统误差太大, $|ZW| > 3$ 说明实验室的随机误差太大^[2]。

3.2 两种稳健统计方法的比较

四分位数和迭代稳健统计方法对数据的分布形态都没有特殊要求, 也无需剔除数据集中的异常值, 因此都具有稳健性, 但数据集的极端数据对四分位数和迭代统计的影响不同。在四分位数稳健统计中, 以中位值作为对总体的估计值, 与平均数相比, 中位数就是相当稳健的, 因为再极端的数据都不会影响到中位数本身的变化。然而中位数也仅仅是基于排序以后的数据, 而且对数据中间部分的舍入和分组也比较敏感, 同时对总体平均水平的度量效率不是很高。而迭代法是以迭代之后数据的平均值作为总体的估计值, 它是通过降低离群值的权重系数, 提高统计量的全局效率和耐抗性, 从而实现其稳健性。

按照式(7), 用两种稳健统计方法计算 CNAS T0497 能力验证计划部分实验室的 Pb 检测结果, 参加实验室的 z 值见表 4。用四分位数和迭代稳健统计计算的 Pb 浓度值分别为 1.095 和 1.091 mg/L, 稳健标准偏差分别为 0.039 和 0.057 mg/L。两种方法对总体的估计一致, 迭代法得出的标准偏差较四分位数法大。迭代稳健统计法计算的 z 比分数值偏离 0 的程度总体上较四分位数法小, 也就是说, 数据分散度较高时, 用迭代稳健统计法进行数据处理, 实验室的检测结果更趋于满意。数据越集中, 两种统计方法的结果越接近。相对于四分位数统计方法, 迭代稳健统计是一种折中的方法, 更加温和可靠。

表 4 CNAS T0497 能力验证计划部分实验室 Pb 的 z 比分数

代码	Pb 浓度	z 比分数 1*	z 比分数 2**
01	1.08	-0.39	-0.18
02	1.07	-0.64	-0.36
03	1.02	-1.93	-1.23
04	1.13	0.90	0.69
05	1.20	2.70	1.91
06	1.07	-0.64	-0.36
07	1.09	-0.13	-0.01
08	1.15	1.41	1.04
09	1.02	-1.93	-1.23
10	1.10	0.13	0.17
11	1.13	0.90	0.69
12	1.10	0.13	0.17
13	1.20	2.70	1.91
14	1.12	0.64	0.51
14	1.12	0.64	0.51
15	1.16	1.67	1.21
16	0.93	-4.24	-2.80
17	1.08	-0.39	-0.18
18	1.10	0.13	0.17
19	1.11	0.39	0.34
20	0.988	-2.75	-1.79
21	1.09	-0.13	-0.01
22	1.07	-0.64	-0.36
23	1.02	-1.93	-1.23
24	1.10	0.13	0.17

注: * 表示用四分位数稳健统计法计算的 z 值, ** 表示用迭代稳健统计法计算的 z 值。

4 结论

四分位数统计方法和迭代统计方法都属于稳健的数据统计方法, 对数据分布没有特殊要求, 不需要剔除数据中的离群值。四分位数统计方法是以中位值作为对总体的估计, 而迭代法则是以迭代后数据的均值作为对总体的估计。数据越集中, 两种统计方法的结果越接近。对 z 值分析表明, 相对于四分位数法, 迭代稳健统计方法是一种更加温和可靠的统计方法。但是对于采用分割水平设计的能力验证, 迭代法不能使用。

参考文献:

- [1] CNAS/RL07: 能力验证规则。
- [2] 吴忠祥. 实验室能力验证中的分割水平检测样品与稳健统计技术[J]. 中国环境监测, 2003, 19(04): 8-10.
- [3] 郭亚帆. 稳健统计以及几种统计量的稳健性比较分析[J]. 统计研究, 2007, 24(09): 82-85.
- [4] ISO13528: 2005, "Statistical Methods for use in Proficiency Testing by Interlaboratory Comparisons", Geneva, Switzerland.