

实验室间比对数据统计分析方法的探讨

济南兰光机电技术有限公司 刘开凤 郝文静 周伟芳

摘要：目前，常用的实验室间比对数据统计分析方法有稳健（Robust）统计法、格拉布斯检验法（Grubbs）及科克伦(Cochran)检验法。基于济南兰光组织的 2014 年春季塑料薄膜阻隔性实验室间比对活动，本文对稳健统计法、格拉布斯检验法及科克伦检验法进行分析比较，探讨了三种数据评价方法在计算过程、离群值判断等方面的区别。

关键词：阻隔性，实验室间比对，稳健统计法，格拉布斯检验法，科克伦检验法

近期，Labthink 兰光 2014 年春季塑料薄膜阻隔性实验室间比对活动已圆满落幕，本次实验室间比对活动采用 CNAS-GL02《能力验证结果的统计处理和评价指南》中推荐的稳健（Robust）统计法对数据进行统计分析评价。而在国外，实验室间比对常用的数据统计分析方法为 ISO 5725.2-1994《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与再现性的基本方法》中推荐的格拉布斯检验法（Grubbs）和科克伦(Cochran)检验法。三种统计方法在计算过程特别是离群值判断方面存在较大差异。基于本次实验室间比对活动，笔者对三种统计方法进行详细介绍与分析。

1. 数据统计方法简介

1.1 稳健统计法

采用稳健统计法进行统计分析的前提是假设分析的结果服从正态分布。在开始进行统计分析之前，应确保所采集的数据是正确、合理的，并对数据中的粗大误差和潜在问题进行识别。通常在数据统计过程中需要统计结果数、中位值、标准四分位间距（标准化 IQR）、稳健 CV、最小值、最大值和极差七种统计量。这些统计量的计算是进行实验室结果统计评价的基础。其中最重要的统计量是中位值和标准化 IQR，它们是数据集中和分散的量度，属于稳健统计量，不受数据中离群值的影响。

对实验室检测结果应使用基于稳健总计统计量的 Z 比分数（中位值和标准

化 IQR) 进行评价。对于分割水平对样品 a 和样品 b 应统计计算两个 Z 比分数——实验室间 Z 比分数 (ZB) 和实验室内 Z 比分数 (ZW)。它们分别基于结果对的标准化和 (S) 和标准化差 (D) 来进行计算, 标准化和 (S) 和标准化差 (D) 的计算公式如下:

$$S = \frac{(a+b)}{\sqrt{2}} \dots\dots\dots \text{公式 (1)}$$

$$D = \frac{(a-b)}{\sqrt{2}} \text{ (D 保留正负号)} \dots\dots\dots \text{公式 (2)}$$

通过计算每个实验室的标准化和 (S) 和标准化差 (D), 可以统计出所有参加实验室的 S 和 D 的中位值和标准化 IQR, 再根据公式 (3) 和公式 (4) 即可计算实验室间 Z 比分数 (ZB) 和实验室内 Z 比分数 (ZW)。

$$ZB = \frac{S - \text{中位值}(S)}{\text{标准化IQR}(S)} \dots\dots\dots \text{公式 (3)}$$

$$ZW = \frac{D - \text{中位值}(D)}{\text{标准化IQR}(D)} \dots\dots\dots \text{公式 (4)}$$

实验室结果依据 Z 比分数 (包括 ZB 和 ZW) 进行评定。 $|Z| \geq 3$ 的结果为离群值, $2 < |Z| < 3$ 时为可疑值。正的实验室间离群 (即 $ZB \geq 3$) 表明该样品对的两个结果太高, 负的实验室间离群值 (即 $ZB \leq -3$) 表明其结果太低, 而实验室内离群值 (即 $|ZW| \geq 3$) 则表明其两个结果间的差值太大。

1.2 科克伦检验法

科克伦检验法应用于所有标准差都是在重复性条件下, 且由相同数目 (n) 的测试结果计算得出的情形。实际中由于数据的缺失或剔除, 测试结果数可能不同。所以假定在正常组织的试验中, 每个实验室的测试结果数目不同所造成的影响是有限且可以忽略的, 科克伦检验中所用的 n 可取多数实验室的测试结果数。

该统计方法是对标准差的最大值进行评定, 属于单侧检验。统计过程为先将全组实验室检测结果标准差从小到大排列, 然后根据标准差的最大值计算科克伦检验的统计量, 见公式 (5)。查阅科克伦检验临界值表, 通过比较相同显著水平

下的临界值来判定该最大标准差是否为离群值。如果标准差的最大值是离群值，则将该值舍去后对剩余的标准差重新计算科克伦检验的统计量。

$$C = \frac{s_{\max}^2}{\sum_{i=1}^p s_i^2} \dots\dots\dots \text{公式 (5)}$$

其中， s_i 为实验室的标准差， s_{\max} 是 s_i 中最大值。

离群值的判断：

- a) 当统计量小于等于 5%的临界值时，认为该值是正常值；
- b) 当统计量大于 5%的临界值且小于等于 1%的临界值时，该值为可疑值；
- c) 当统计量大于 1%临界值时，认为该值是统计离群值。

1.3 格拉布斯检验法

格拉布斯检验法适用于判定一组检测结果中最大值或最小值是否为离群值，此方法通过计算平均值 \bar{x} 和标准偏差 SD 然后再计算统计量 (G)，见公式 (6)、公式 (7)。通过查阅 Grubbs 检验临界值表获得一定显著性水平 (α) 下的临界值，比较统计量 (G) 与查到的临界值。当统计量 (G) 大于临界值时，此数据即为离群值，并对剩余数据继续检验，直到数据中无异常值为止。

Grubbs 检验的统计量计算公式为：

$$G_p = (x_p - \bar{x}) / SD \dots\dots\dots \text{公式 (6)}$$

其中， x_p 为一系列数据 x_i ($i=1,2, \dots, p$) 中的最大值； \bar{x} 为平均值；SD 为标准偏差。

最小值 x_1 的统计量计算公式为：

$$G_1 = (\bar{x} - x_1) / SD \dots\dots\dots \text{公式 (7)}$$

其中， x_1 为一系列数据 x_i ($i=1,2, \dots, p$) 中的最小值； \bar{x} 为平均值；

SD 为标准偏差。

离群值的判断：

- a) 当统计量小于等于 5%的临界值时，认为该值是正常值；
- b) 当统计量大于 5%的临界值且小于等于 1%的临界值时，该值为可疑值；

c) 当统计量大于 1%临界值时，认为该值是统计离群值。

2.实验室间比对结果分析

本次塑料薄膜阻隔性实验室间比对活动采用“分割水平”的样品设计，氧气透过量的样品分为 A、B 两组，水蒸气透过量的样品分为 C、D 两组。在对实验室检测能力进行评定时，不出具“合格”与否的结论，而是使用“满意”、“可疑”或“离群”的概念。如果一个结果被评定为离群值，这意味着从统计上看它明显地不同于同组的其他结果。

2.1 采用稳健统计法评定结果

用稳健统计法计算各参加实验室的实验室间 Z 值(ZB)和实验室内 Z 值(ZW)。通过对检测结果进行统计分析，本次氧气透过量检测项目发现实验室间离群 1 家，实验室间可疑 2 家，实验室内离群 0 家，实验室内可疑 6 家；对于水蒸气透过量检测项目，此次有实验室间离群 2 家，实验室间可疑 1 家，实验室内离群 7 家，实验室内可疑 9 家。将实验室 Z 值 (ZB 和 ZW) 合并统计，本次氧气透过量项目共有 94%的实验室结果满意，5%的实验室可疑，1%的实验室离群；水蒸气透过量项目共有 84%的实验室结果满意，8%的实验室可疑，8%的实验室离群。

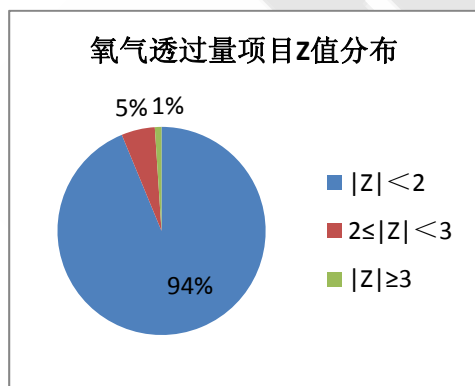


图 1: 氧气透过量项目 Z 值分布

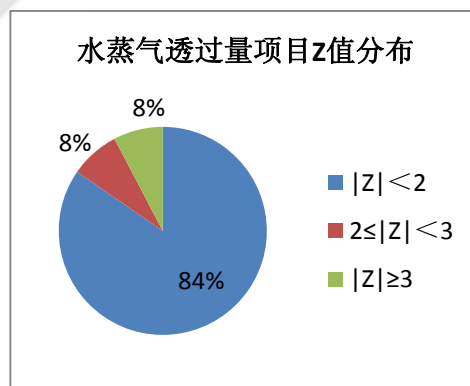


图 2: 水蒸气透过量项目 Z 值分布

由统计数据可见，实验室内的离群与可疑明显多于实验室间离群与可疑，说明实验室内部检测过程中的随机误差比较多，而实验室间的系统误差要少一些。

2.2 采用科克伦检验法评定结果

用科克伦检验法进行数据统计时，需要计算统计量 C ，再与临界值进行比较判定离群值。科克伦检验法是对检测结果的标准偏差进行统计以剔除离群值。经过统计计算，本次实验室比对活动中，样品 A 的可疑值有 1 个，离群值有 9 个；样品 B 有可疑值 1 个，离群值 8 个；样品 C 有可疑值 2 个，离群值 13 个；样品 D 有可疑值 5 个，离群值 12 个。

根据科克伦统计原理，采用此方法时，参与实验室数量越多，其对应的临界值越小，即要求标准偏差的离散程度要小。参与实验室数量少时，对应的临界值大，那么对于数据标准偏差离散程度的要求就不是那么严格。因此，在此次实验室间比对活动参与实验室较多，而数据的离散程度偏大的情况下，科克伦检验法评定剔除的离群值相对较多。

2.3 采用格拉布斯检验法评定结果

在此次实验室间比对中，对同一种样品的检测结果呈两端离散的形式，所以在采用格拉布斯检验法进行数据统计时，需要同时对最大值和最小值计算统计量 G_p 和 G_1 ，再与临界值 $G(\alpha)$ 比较。如果 $G_p > G_1$ 且 $G_p > G(\alpha)$ ，则评定最大值为离群值；如果 $G_1 > G_p$ 且 $G_1 > G(\alpha)$ ，则评定最小值为离群值。对于样品 A，经计算没有发现可疑值和离群值；对于样品 B，经计算发现可疑值 2 个，离群值 1 个；对于样品 C，计算发现可疑值 2 个，离群值 2 个；对于样品 D，计算没有发现可疑值和离群值。

格拉布斯检验法是对实验室检测结果的平均值进行统计，从而剔除离群值。在数据量较少时，其对应的临界值越小，那么剔除离群值的概率就较大。在数据量较大时，其对应的临界值越大，那么剔除离群值的概率就较低。因此，在此次实验室比对活动中，采用该方法计算时发现的离群值较少。

2.4 不同统计方法实验室比对结果评定的比较

对于本次实验室间比对样品的参考值，不同的统计方法依据不同的统计量来反映。稳健统计法采用中位值反映样品的参考值，而科克伦检验与格拉布斯检验则采用总平均值。总平均值并不是对一组数据计算其平均值得到的，而是首先通过科克伦检验法对标准偏差进行评价剔除离群值，然后通过格拉布斯检验法剔除平均值的离群值，之后再对剩余的数据计算平均值即为总平均值。

通过对本次实验室间比对数据的计算，得到的参考值与离群情况见表 1、表 2:

表 1: 实验室间比对参考值比较

统计方法	样品 A ($\text{cm}^3/\text{m}^2 \cdot 24\text{h} \cdot 0.1\text{MPa}$) 或 $\text{ml}/\text{m}^2 \cdot \text{day}$	样品 B ($\text{cm}^3/\text{m}^2 \cdot 24\text{h} \cdot 0.1\text{MPa}$) 或 $\text{ml}/\text{m}^2 \cdot \text{day}$	样品 C (g/ $\text{m}^2 \cdot 24\text{h}$)	样品 D (g/ $\text{m}^2 \cdot 24\text{h}$)
稳健统计法	58.429	17.611	5.433	2.859
科克伦 检验+格 拉布斯 检验	59.770	17.879	5.368	2.987

表 2: 实验室间比对结果统计

统计方法	样品	满意 (个)	可疑 (个)	离群 (个)	备注	
稳健统计法	ZB	样品 A+B	101	2	1	样品 A、B 用于 氧气透过量测 试; 样品 C、D 用 于水蒸气透过 量测试。
		样品 C+D	108	1	2	
	ZW	样品 A+B	98	6	0	
		样品 C+D	95	9	7	
科克伦检验+格 拉布斯检验	样品 A	95	1	9		
	样品 B	96	3	9		
	样品 C	96	4	15		
	样品 D	99	5	12		

由表 1、表 2 可以看出，虽然通过稳健统计法判定出的可疑值与离群值数量与通过科克伦检验法和格拉布斯检验法剔除的可疑值与离群值数量有较大的差别，但是两种统计方法计算的样品参考值是很接近的。由此可以看出，稳健统计法和由科克伦检验法与格拉布斯检验法组合检验这两种检验方法的统计参数受极端值的影响都不大。

3. 总结

通过此次实验室间比对数据的分析统计，可以看出，稳健统计法不受数据量多少的影响，而且受极端值的影响也不大。而由科克伦检验法与格拉布斯检验法组合检验的统计方法，受数据量多少的影响较大，而受极端值的影响较小。不过两种统计方法相比较，稳健统计法在数据处理和计算程序上更加科学、直观和简洁方便。