

## 第十六章 实验方法评价与数据处理

### 第一节 实验方法评价

随着食品科学的不断发展，食品检验方法不断更新，评价检验方法的标准也逐步建立和完善起来。这些评价标准主要是准确度、精密度、检测限以及费用与效益。

#### 一. 评价指标

##### 1、准确度

准确度是指在一定条件下，多次测定的平均值与真实值相符合的程度。准确度通常用绝对误差或相对误差表示。

在实际工作中，一般在试样中添加已知标准物质质量作为真值，并以回收率表示准确度。即：

$$P(\%) = \frac{X_1 - X_0}{m} \times 100\% \quad 16-1$$

式中：P (%) ——加入标准物质的回收率；

$X_1$  ——加标样品测定值

$X_0$  ——试样本底测定值

$m$  ——加入标准物质的质量

式中的本底值  $X_0$ ，其测定精密度所显示误差是反映随机误差；加入标准物质的质量  $m$ ，其测定误差反映了系统误差。所以，回收率是两种误差综合指标，能决定方法的可靠性。对回收率的要求数值是个比较复杂的问题，依分析测定方法难易和不同类型的分析方法而变化。一般 ppm 级应在 90% 以上；ppb 级如荧光法测定苯并芘在 80%；比较繁杂的方法 70% 即可；但最低不能小于 70%。

##### 2、精密度

精密度是指多次重复测定某一样品时，所得测定值的离散程度。精密度通常用标准差或相对标准差来表示。

重复测定的精密度与待测物质绝对量有关，一般规定：mg 级  $C_v$  (变异系数或相对标准差) 为 5%； $\mu\text{g}$  级  $C_v$  为 10%，ng 级  $C_v$  为 50% 左右。

##### 3、检测限

检测限是指分析方法在适当的置信水平内，能从样品检测被测组分的最小量或最小浓度，即断定样品中被测组分的量或浓度确实高于空白中被测组分的最低量。

一般对检测限有几种规定方法：

(1) 气相色谱法：用最小检测量或最小检测浓度表示

最小检测量是指检测器恰能产生色谱峰高大于二倍噪音时的最小进样量。即：

$$S = 2N \quad 16-2$$

式中：S --- 最小响应值

N --- 噪音信号

最小检测浓度是指最小检测量与进样量体积之比。即单位进样量相当待测物质的量。

**例 16-1** 用气相色谱法测定聚氯乙烯成型品中氯乙烯单体的检测限。仪器噪音的最大信号为峰高 1.0mm，注入 0.5 $\mu\text{g}$  氯乙烯标准制备的顶空气 3ml，响应值为 12mm，求最低检测量？

题解：根据公式[16-2]  $S = 2N = 2 \times 1.0\text{mm} = 2.0\text{mm}$

由

$$\frac{\text{最小检测量}}{\text{最小响应值}} = \frac{\text{注入标准物质量}}{\text{标准物质的响应值}} \quad \text{可知道}$$

$$\text{最小检测量} = \frac{\text{最小响应值} \times \text{注入标准物质量}}{\text{标准物质的响应值}} = 2 \times 0.5 / 12 = 0.083 \mu\text{g}$$

**例 16--2** 在上例中, 如果注入 3ml 顶空气相当于 0.5g 聚氯乙烯成型品, 求最小检测浓度?

题解: 最小检测浓度 = 0.083 / 0.5 = 0.17 μg/g

### (2) 吸光光度法

在吸光光度法中, 扣除空白值后, 吸光值为 0.01 所对应的浓度做为检测限。

**例 16-3** 利用镉离子与 6-溴苯并噻唑偶氮萘酚形成红色络合物, 对食品中镉含量进行比色测定。对全试剂空白进行 5 次平行测定, 吸光度平均值是 0.003, 再测定 0.25 μg 标准镉溶液, 其吸光度为 0.023, 求检测限?

题解: 由  $\frac{\text{检测限}}{\text{最小响应值}} = \frac{\text{镉标准质量}}{\text{镉标准吸光度} - \text{空白吸光度}}$  可以知道:

$$\text{检测限} = \frac{\text{最小响应值} \times \text{镉标准质量}}{\text{镉标准吸光度} - \text{空白吸光度}} = 0.01 \times 0.25 / (0.023 - 0.003) = 0.125 \mu\text{g}$$

(3) 一般实验: 当空白测定次数  $N > 20$  时, 给出置信水平 95%, 检测限为空白值正标准差 (S) 的 4.6 倍。即: 检测限 = 4.6 × S 16--3

**例 16-4** 在例 16-3 中, 当空白测定  $N > 20$  时, 吸光度  $0.003 \pm 0.001$  相当于镉  $0.0375 \pm 0.013 \mu\text{g}$  求检测限?

题解: 按式 [16-3] 检测限 = 4.6 × S = 4.6 × 0.013 = 0.06 μg

若空白测定次数  $n < 20$  时, 检测限按下式计算

$$\text{检测限} = 2\sqrt{2} t_f S \quad 16--4$$

式中:  $t_f$  置信水平为 95% (单侧), 批内自由度为  $f$  时的临界值

$f$  - 批内自由度

$f = m(n-1)$ ,  $m$  为重复测定次数,  $n$  为平行测定次数。

**例 16-5** 用 2, 3-二氨基萘荧光法测定硒, 双空白测定 10 次, 其空白值为  $11.4 \pm 1.3 \text{ng}$ , 求检测限?

题解: 根据  $f = m(n-1) = 10 \times (2-1) = 10$

查附表, 当置信水平 95% (单侧) 时,  $t_{10} = 2.23$

$$\text{按式 [16-4], 检测限} = 2\sqrt{2} t_f S = 2 \times 1.414 \times 2.23 \times 1.3 = 8.20 \text{ng}$$

### (4) 国际理论应用化学联合会对检测限的规定

对于各种光学分析方法, 可测量的最小分析响应值以下式表示

$$X_L = \bar{X}_b - K \times S_b \quad 16--5$$

式中:  $X_L$  - 最小响应值;

$\bar{X}_b$  - 多次测量空白值的平均值 ( $n \geq 20$ );

$S_b$  - 多次测量空白值的标准差;

K--根据一定置信水平确定的系数（一般当置信水平为 90%，空白测量次数  $n < 20$  时， $K=3$ ；置信水平为 95%， $n > 20$  时， $K=4.65$ ）。

$$\text{检测限} = \frac{X_L - \bar{X}_b}{m} = \frac{K \times S_b}{m} \quad 16-6$$

式中， $m$  为方法灵敏度，即单位浓度或单位量被测物质所产生的响应值的变化程度，在实际工作中，以标准曲线斜率度量灵敏度，其余各项符号同式[16-5]。

**例 16-6** 在例 16-5 测定硒时，增加空白测定数，其空白值为： $10.1 \pm 0.95\text{ng}$ ，其灵敏度为 0.54 荧光单位/ng，求检测限？

题解：按 16-6 式

$$\text{检测限} = \frac{X_L - \bar{X}_b}{m} = \frac{K \times S_b}{m} = 3 \times 0.95 / 0.54 = 5.28\text{ng}$$

从检测限定义可以知道，增加实际测定次数，提高测定精密度，降低仪器噪音，可以改善检测限。

#### 4、费用与效益

费用与效益是目前国内外重视的问题。实验室工作人员，应结合实际测试目标，选择或设计相应准确度和精密度的实验方法。用一般常规实验能够完成的测定，不必使用贵重精密仪器。检验员经训练能较好掌握某种测定方法的时间，也是评价实验方法的重要内容。“简单易学”在一定程度上意味着能保证检验质量。从实际工作需要出发，快速，微量，费用低廉，技术要求不高，操作安全的测定方法应列为一般实验室的首选方法。

#### 二、实验结果的检验

在食品分析中，常遇到两个平均值的比较问题，如测定平均值和已知值的比较，不同分析人员，不同实验室，或用不同分析方法测定的平均值的比较，对比性试验研究等均属于此类问题。所以对这类问题常采用显著性检验法-----利用统计方法来检验被处理问题是否存在统计上的显著性，常用有  $t$  检验法和  $F$  检验法。

##### 1、 $t$ 检验法

用以比较一个平均值与标准值之间或两个平均值之间是否有在显著性差异。进行  $t$  检验的程序如下：

(1) 选定所用的检验统计量：

当检验样本均值  $\bar{X}$  与总体均值  $\mu$  是否有显著性差异时，使用统计量

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}} \quad 16-7$$

式中： $S$ —标准差

当检验两个均值之间是否有显著性差异时，使用统计量

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\bar{S}} \times \sqrt{\frac{n_1 \times n_2}{n_1 + n_2}} \quad 16-8$$

其中  $\bar{S}$  --合并标准差 按下式计算：

$$\bar{S} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \quad 16-9$$

式中  $S_1^2$  ---第一个样本的方差

$S_2^2$  ---第二个样本的方差

$n_1$ ---第一个样本的测定次数

$n_2$ ---第二个样本的测定次数

(2) 计算统计量,

如果由样本值计算的统计量值大于 t 分布表中相应显著性水平  $\alpha$  和相应自由度 f 下的临界值  $t_{\alpha, f}$ , 则表明 被检验的均值有显著性差异,反之, 差异不显著。

应用 t 检验时, 要求被检验的两组数据具有相同或相近的方差(标准差)。因此在 t 检验之前必须进行 F 检验, 只有在两方差一致性前提下才能进行 t 检验。

(3) 假设检验的一尾测验与两尾测验

在进行测验结果分析确定检验水平时, 还应根据其处理的性质和试验结果的准确性, 考虑显著性测验用一尾测验还是用两尾测验。

在提出一个统计假设时, 必然有一个与其相对应的备择假设。备择假设为否定假设时, 必然接受的另一个假设。例如, 单个平均数进行显著性测验时, 通常  $H_0: \mu=\mu_0$ ,  $H_A: \mu\neq\mu_0$ 。如果  $H_0$  被否定接受  $H_A$  时, 其  $\mu\neq\mu_0$ , 便有  $\mu$  大于  $\mu_0$  或小于  $\mu_0$  的两种可能性, 即所测定的误差概率在正态分布曲线的左尾和右尾各有一个否定域, 而附表\_1\_\_中临界 t 值表规定的  $\alpha$  值是两尾概率之和。如果确定的检验水平  $\alpha=0.05$ , 则两尾否定域的概率各为 0.025, 这类测验称这为两尾测验。

但有的试验则不然, 例如某酿酒厂曲种酿造醋的醋酸含量大于 12%, 则其假设  $H_0: \mu > 12\%$ ,  $H_A: \mu \leq 12\%$ 。如果选择的曲种酿造醋酸含量大于 12%,  $H_0$  被否定,  $\mu$  只能大于 12%。若小于 12%便不符合规定的企业标准, 没有推广价值, 因此只有在正态曲线的右尾一个否定域, 这类检验称为一尾测验。两尾测验查两尾概率表或一尾测验查一尾概率表时, 可以直接从表上查得。如果两尾检验查一尾概率表时, 需将检验水平值除 2, 再查出  $\mu_\alpha$  值, 如两尾测验检验水平  $\alpha=0.05$ , 一尾概率  $\mu_{0.05}=1.64$ , 应将检验水平  $\alpha=0.05$  除以 2 得  $\alpha=0.025$ ,  $\mu_{0.025}=2.24$ , 如果一尾检验查两尾概率表时需将检验水平值乘以 2, 再查出  $\mu_\alpha$  值, 如一尾检验水平  $\alpha=0.1$ ,  $\mu_{0.1}=1.64$ , 因此用一尾测验还是用两尾测验, 应认真从实际考虑。

而 t 检验法为判别性测验, 多为两尾测验。

下面将 t 检验法在食品分析中的主要应用介绍如下:

①用已知组成的标样评价分析方法

为了鉴定一个分析方法的可靠性, 可用一已知量的基准物或已知含量的标准试样进行对照试验, 通过若干次测定, 取得其平均值, 然后将这个平均值与已知值(真值)进行比较, 从而判断这个分析方法是否存在系统误差。因为这时将平均值与真值进行比较, 所以可以按 t 检验法来判别。逻辑推理是先假设平均值与真值之间不存在真正的差异, 如果所算出的 t 值大于通常规定的置信水平的 t 值, 那么, 应该拒绝所提的假设, 就是说, 这样的差异不能认为是偶然的误差, 而是被检验的方法存在系统误差, 反之, 则应接该受假设, 判断该方法不存在系统误差。

**例 16-7** 为了鉴定一个分析方法的准确度, 取重量为 100mg 的基准物进行 10 次测定, 所得数据为 100.3, 99.2, 99.4, 100.0, 99.7, 99.9, 99.4, 100.1, 99.4, 99.6, 试对这组数据进行评价。

题解: 计算平均值和标准偏差。  $\bar{X}=99.7$ ,  $S=0.38$ , 按式 16—7, 计算统计量

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{99.7 - 100}{0.38/\sqrt{10}} = -2.50 \quad \text{查 } t \text{ 表得 } t_{0.05,10}=2.23, \quad |t| > t_{0.05,10}=2.23$$

表明 10 次测定的平均值与标准值有显著性差异，可认为该方法存在系统误差。

### ②两个平均值的比较

在进行分析方法研究的时候，往往要在两种分析方法之间，两个不同实验室之间或两个不同操作者之间进行比较试验。这时对同一试样各测定若干次，得到两组测定数据的平均值，以比较两个平均值来判断它们之间是否存在真正的差异。如果两组测定数据的精密度高，两个平均值相差又比较大，这种情况自然容易判断。有时，两组数据本身不很精密，而两个平均值相差又不太大，这就是利用统计分析法才能进行正确判断。

两组测定的平均值都不是真值，在进行检验时，将两组数据看作同属一个总体来处理，按式 16-8, 16-9 计算统计量  $t$ ，与查表所得的  $t$  值 ( $f=n_1+n_2-2$ ) 进行比较，便能作出判断。

**例 16-8** 采用两种不同方法测定奶粉中脂肪含量，测定数据见表 16-1。试比较两种方法的精密度有无显著差异。

**表 16-1 奶粉脂肪测定数据%**

方法 1			方法 2		
脂肪含量	$ X_{1i} - \bar{X}_1 $	$ X_{1i} - \bar{X}_1 ^2$	脂肪含量	$ X_{2i} - \bar{X}_2 $	$ X_{2i} - \bar{X}_2 ^2$
2.01	0.04	0.0016	1.88	0.04	0.0016
2.10	0.13	0.0169	1.92	0.00	0.0000
1.86	0.11	0.0121	1.90	0.02	0.0004
1.92	0.05	0.0025	1.97	0.05	0.0025
1.94	0.03	0.0009	1.94	0.02	0.0004
1.99	0.02	0.0004			
$\sum$ 11.82		0.0344	9.61		0.0048

题解：根据两组数据，分别计算两种方法得平均值  $\bar{X}$  及标准差  $S$ 。

$$\bar{X}_1 = \frac{\sum X_{1i}}{n} = 1.97 \quad S_1 = \sqrt{\frac{\sum (X_{1i} - \bar{X}_1)^2}{n-1}} = 0.083$$

$$\bar{X}_2 = \frac{\sum X_{2i}}{n} = 1.92 \quad S_2 = \sqrt{\frac{\sum (X_{2i} - \bar{X}_2)^2}{n-1}} = 0.035$$

计算合并标准差

$$\bar{S} = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} = \sqrt{\frac{5 \times 0.083^2 + 4 \times 0.035^2}{6+5-2}} = 0.066$$

计算统计量

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\bar{S}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{1.97 - 1.92}{0.066} \sqrt{\frac{5 \times 6}{5 + 6}} = 1.383$$

查 t 表得  $t_{0.05,9} = 2.2622$ ,  $t < t_{0.05,9}$ , 两法差别不显著, 即两种测定结果是一致的。

### ③ 配对比较试验数据

在分析方法试验中, 为了判断某一个因素的结果否有显著影响, 往往取若干批的试样, 将其他因素固定下来, 对某一因素进行配对的比较试验。这样的试验可以消除其它因素的影响而把被检验的因素突出出来, 以便从随机误差的覆盖下找出被检验的因素是否存在真正的差异。例如为了比较两个实验室的分析结果, 取若干批试样交由两个实验室进行比较测定; 为了比较两种分析方法的差异性, 可以用两种不同方法对同一试样进行测定比较也可以把一个试样交给几个人进行方法的比较试验等等。

配对比较试验数据的判断, 不是根据两组数据的平均值来进行比较, 而是根据各组配对数据之差 D 来进行显著性的检验。

首先计算配对数据之差 D 的平均值  $\bar{D}$ , 标准差  $S_D$

$$\bar{D} = \frac{\sum D_i}{n} \quad 16-10$$

$$S_D = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{\sum D_i^2 - (\sum D_i)^2/n}{n-1}} \quad 16-11$$

然后计算统计量 t

$$t_D = \frac{\bar{D} \times \sqrt{n}}{S_D} \quad 16-12$$

如果计算的统计量值小于 t 分布表中相不在显著性水平  $\alpha$  和相应自由度 f 的临界值  $t_{\alpha, f}$ , 则表明被检验的两种方法测定结果是一致的。

**例 16-9** 某实验室使用直接离子计测定饮料中的氟含量。为了检验新方法的可靠性, 用新法和老法 (氟试剂比色法) 同时对 10 份不同饮料进行了对比性测定, 结果见表 16-2, 两法的测定结果是否一致?

**表 16-2 两种方法测定氟含量数据**

饮料样品	氟含量 (mg/L)		相差值 $D_i$	$D_i^2$
	氟试剂比色法	直接离子计法		
1	4.18	4.42	-0.24	0.0576
2	4.04	4.17	-0.13	0.0169
3	4.36	3.14	1.22	1.4884
4	3.01	2.94	0.07	0.0049
5	1.66	1.20	0.46	0.2116
6	10.31	7.96	2.35	5.5225
7	5.92	9.80	-3.88	15.0544
8	2.5	1.43	1.07	1.1449
9	5.98	3.97	2.01	4.0401
10	6.56	4.83	1.73	2.9929

Σ			4.66	30.5342
---	--	--	------	---------

题解：按式 16—10，16—11 计算差数的平均值  $\bar{D}$  与标准差  $S_D$

$$\bar{D} = \frac{\sum D_i}{n} = 4.66/10 = 0.466 \quad \text{mg/L}$$

$$S_D = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{\sum D_i^2 - (\sum D_i)^2/n}{n-1}} = 1.78 \text{mg/L}$$

$$t_D = \frac{\bar{D} \times \sqrt{n}}{S_D} = \frac{0.466 \times \sqrt{10}}{1.78} = 0.83$$

按式 16—12 计算 t 值

查 t 分布表，当  $f=9$  时， $t_{0.05,9}=2.26$ ， $t < t_{0.05,9}=2.26$ ，说明测定结果的差别无显著性，即两种方法的测定结果是一致的。

## 2、F 检验法

F 检验法是通过计算两组数据的方差之比来检验两组数据是否存在显著性差异。比如使用不同的分析对同一试样进行测定得到的标准差不同，或几个实验室用同一种分析方法测定同一试样，得到的标准差不同，这时就有必要研究产生这种差异的原因，通过这种 F 检验法，可以得到满意的解决。

F 检验法其步骤如下：

(1) 计算统计量方差比

$$F = \frac{S_1^2}{S_2^2}$$

16—13

式中  $S_1^2, S_2^2$  ----- 分别代表两组测定值的方差

(2) 查 F 分布表

(3) 判断：当计算所得 F 值大于 F 分布表中相应显著性水平  $\alpha$  和自由度  $f_1, f_2$  下的临界值  $f_{\alpha}(f_1, f_2)$ ，即  $F > f_{\alpha}(f_1, f_2)$  时，则两组方差之间有显著性差异，反之，则两组方差无显著性差异。

在编制 F 分布表时，是将大方差作分子，小方差作分母，所以，在由样本值计算统计量 F 值时，也要将样本方差  $S_1^2, S_2^2$  中数值较大的一个作分子，较小的一个作分母。

**例 16-10** 仍以例 16-8 中实验数据为例，通过 F 检验法比较两种方法的精密度有无显著差异。

题解：分别计算两种方法的方差

$$S_1^2 = 0.083^2 = 0.0069$$

$$S_2^2 = 0.035^2 = 0.0012$$

按式 16-13 计算统计量方差比 F：

$$F = \frac{S_1^2}{S_2^2} = 0.0069/0.0012 = 5.75$$

查 F 分布表,  $F_{0.05}(5, 4) = 6.26$ ,  $F < F_{0.05}(5, 4) = 6.26$

说明差别不明显, 即两种测定方法精密度是一致的。

**例 16-11** 用原子吸收法与比色法同时测定某试样中的铜, 各进行了 10 次测定, 原子吸收法测定方差为  $6.5 \times 10^{-4}$ , 比色法测定的方差为  $8.0 \times 10^{-4}$ , 试由测定精密度考虑, 以选取那一测定方法合适。

题解: (1) 给定量著性水平  $\alpha = 0.10$ , 根据本例题意, 只要检验两个方差是否有显著性差异, 不管两个方差中哪一个比另一个大得多或小得多, 都认为是有显著性差异, 因此是双侧检验。F 分布表中给示的单侧检验 F 临界值, 对于双尾检验, 在给定显著性水平  $\alpha$  时, 要从 F 分布表中查  $F_{\alpha/2}$  值, 针对本例情况, 查 F 分布表,  $F_{0.05(9,9)} = 3.18$

(2) 计算统计量得

$$F = \frac{S_1^2}{S_2^2} = 8.0 \times 10^{-4} / 6.5 \times 10^{-4} = 1.23$$

$F < F_{0.05(9,9)} = 3.08$  说明不能认为两法方差有显著性差异, 即选用原子吸收或比色法都是可以的。

## 第二节 实验数据处理

### 一、分析结果的表示

食品分析项目众多, 某些项目测验结果还可以用多种化学型式来表示, 如硫含量, 可用  $S^{2-}$ ,  $SO_2$ ,  $SO_3$ ,  $SO_4^{2-}$  化学形式表示, 它们的数值各不相同。测定结果的单位也有多种形式, 如 mg/L, g/L, mg/kg, g/kg, mg/100g, 百分含量 (%), 等等, 取不同单位时显然结果的数值不同。

统计处理结果的表示方法也多种多样, 如算术平均值  $\bar{X}$ , 极差, 标准偏差等表示测定数据的离散程度 (精密度)。

原则上讲, 食品分析要求提出的测定结果既反映数据的集中趋热, 又反映测定精密度及测定次数, 另外还要照顾食品分析自身的习惯表示法。

通常, 食品的分析中报出的测定结果的单位采用 mg% 或百分含量%, 而对食品中微量无素的测定结果采用 mg/kg 即 ppm 或  $\mu\text{g}/\text{mg}(\text{ppb})$ , 统计处理的结果采用测定值的算术平均数  $\bar{X}$  与相差  $R = X_{\text{max}} - X_{\text{min}}$  同时表示。当测定数据的重现性较好时, 测定次数 n 通常为 2 次, 当测定数据的重视性较差时, 分析次数应相应地增加。

### 二、实验数据的处理

通过测定工作获得一系列有关分析数据后, 需按以下原则记录, 运算与处理。

#### (一) 记录与运算规则

食品分析中数据记录与计算均按有效数字算法进行, 即:

- (1) 除有特殊规定外, 一般可疑数为最后一位, 有  $\pm 1$  个单位的误差。
- (2) 复杂运算时, 其中间过程可多保留一位, 最后结果须取应有的位数。
- (3) 加减法计算结果, 其小数点以后保留的位数, 应与参加运算各数中小数点以后位数最少者相同。
- (4) 乘法除法计算结果, 其有效数字保留的位数, 应与参加运算各数中有效数字位数最少者相同。

## (二) 可疑数据的检验与取舍

### 1. 实验中的可疑值

在实际分析测试中, 由于随机误差的存在, 使得多次重复测定的数据不可能完全一致, 而存在一定的离散性, 并且常常发现一组测定使其中某一、两个测定值比其余测定值明显的偏大或偏小, 这样的测定值称为可疑值。

可疑值可能是测定值随机流动的极度表现。它虽然明显偏离其余测定值, 但仍然是处于统计上所允许的合理误差之内, 与其余测定值属于同一总体称之为极值, 极值是一个好值, 必须保留, 然而也有可能存在这样的情况, 就是可疑值与其余测定值并不属于同一总体, 称其为界外值、异常值、坏值, 应淘汰不要。

对于可疑值, 必须首先从技术上设法弄清楚其出现的原因。如果查明是由实验技术上的失误引起的, 不管这样的测定值是否为异常值都应舍弃, 而不必进行统计检验。但是, 有时由于各种缘故未必能从技术上找出它出现过失的原因, 在这种情况下, 既不能轻易地保留它, 也不能随意地舍弃它, 应对它进行统计检验, 以便从统计上判明可疑值是否为异常值。如果一旦确定为异常值, 就应从这组测定中将其除掉。

### 2. 舍弃异常值的依据

对于可疑值究竟是极值还是异常值的检验, 实质上就是区分随机误差和过失误差的问题。因为随机误差遵从正态分布的统计规律, 在一组测定值中出现大偏差的概率很小的。单次测定值出现  $\mu \pm 2\sigma$  ( $\sigma$  为标准差, 也用  $S$  表示) 之间的概率为 95.5% (这一概率也称为置信概率或置信度  $\mu \pm 2\sigma$  将为置信区间), 也就是说偏差  $> 2\sigma$  的出现概率为 5% (这概率也称为显著概率或显著性水平); 而偏差  $> 3\sigma$  的概率更小, 只有 0.3%。通常分析检验只进行少数几次测定, 按常规来说, 出现大偏差测定值的可能性理应是极小的, 而现在竟然出现了, 那么就有理由将偏差很大的测定值作为与其余的测定值来源于不同的总体异常值舍弃它。并将  $2\sigma$  和  $3\sigma$  称为允许合理误差范围, 也称为临界值。

### 3. 可疑值的检验准则

①已知标准差 如果人们在长期实践中已知道了标准差  $\sigma$  的数值, 可直接用  $2\sigma$  (置信度 95.5%) 或  $3\sigma$  (置信度 99.7) 作为取舍依据

②未知标准差 一般情况下, 总体标准差  $\sigma$  事先并不知道, 而要由测定值本身来计算它, 并依次来检验该组测定值中是否混有异常值, 判别方法有许多, 如狄克逊 (Dixon) 检验法, 格鲁布斯 (Grubbs) 检验法、科克伦 (Cochran) 最大方差检验法等。下面介绍前两种方法:

#### (1) 狄克逊 (Dixon) 检验法

此法也叫  $Q$  统计量法, 是指用狄克逊法检验测定值 (或平均值) 的可疑值和界外值的统计量, 并以此来决定最大或最小的测定值 (或平均值) 的取舍。其中提到关于平均值的取舍问题, 是由于有时要进行几组数据的重复测定, 取几次测定值的平均值, 也有一个可疑值取舍问题, 也要进行检验。

现将  $Q$  统计量法和检验步骤和方法说明如下:

①首先将一组测定值按大次序排列: 即:  $X_1 \leq X_2 \leq X_3 \dots \leq X_{n-1} \leq X_n$  不言而喻, 异常值 (界外值) 必然出现在两端。

②用表 16-3, 所列公式, 计算  $Q$  统计量。计算时,  $Q$  统计量的有效数字应保留至小数点后 3 位。

③从表 16-4 查出检验显著概率为 5% 和 1% 的  $Q$  统计量的临界值  $Q_{0.05, (H)}$  和  $Q_{0.01, (H)}$ , 其中  $H$  为受检验的一组按小到大排列的测定值的最大的一个序数 (也就是测定次数), 从受检验的测定值的两个  $Q$  统计量计算值中, 只选取较大的  $Q$  统计量的计算值与  $Q$  统计量的临界值比较。

④判定：若计算统计量  $Q \leq Q_{0.05(H)}$ ，则受检验的测定值正常接受。

若  $Q_{0.05(H)} \leq H \leq Q_{0.01(H)}$ ，则受检验的测定值为可疑值。用 1 个星号 "\*" 记在右上角。查有技术原因的可疑值舍去，否则保留。

若  $Q > Q_{0.01(H)}$  则受检验的测定值判为界外值（异常值），用 2 个星号 "\*" 记在右上角。该值舍去。

⑤当 Z (1) 或 Z (H) 舍去时，还需对 Z (2) 或 (H-1) 再检验，注意此时统计量的临界值应为  $Q_{0.05(H-1)}$  和  $Q_{0.01(H-1)}$  依次类推。但在舍去第二个测定值时要慎重考虑有否其他原因。

表 16-3 统计计算方式

H	计算公式	公式用途
3--7	$Q_{10} = \frac{Z(2) - Z(1)}{Z(H) - Z(1)}$ 16--14	检验最小值 Z (1)
	$Q_{10} = \frac{Z(H) - Z(H-1)}{Z(H) - Z(1)}$ 16--15	检验最大值 Z (H)
8--12	$Q_{11} = \frac{Z(2) - Z(1)}{Z(H-1) - Z(1)}$ 16--16	检验最小值 Z (1)
	$Q_{11} = \frac{Z(H) - Z(H-1)}{Z(H) - Z(2)}$ 16--16	检验最大值 Z (H)
13 个以上	$Q_{22} = \frac{Z(3) - Z(1)}{Z(H-2) - Z(1)}$ 16--17	检验最小值 Z (1)
	$Q_{22} = \frac{Z(H) - Z(H-2)}{Z(H) - Z(3)}$ 16--18	检验最大值 Z (H)

用狄克逊检验准则检验的优点是方法简便，概率意义明确，现以气相色谱分析的一个实例来说明具体检验方法。

表 16-4 狄克逊法界外值检验的临界值

临界值			临界值		
H	5%	1%	H	5%	1%
3	0.970	0.994	22	0.468	0.544
4	0.829	0.926	23	0.459	0.535
5	0.710	0.821	24	0.451	0.526
6	0.628	0.740	25	0.443	0.517
7	0.569	0.680	26	0.436	0.510

8	0.608	0.717	27	0.429	0.502
9	0.564	0.672	28	0.423	0.495
10	0.530	0.635	29	0.417	0.489
11	0.502	0.605	30	0.412	0.483
12	0.479	0.579	31	0.407	0.477
13	0.611	0.697	32	0.402	0.472
14	0.586	0.670	33	0.397	0.467
15	0.565	0.647	34	0.393	0.462
16	0.546	0.627	35	0.388	0.458
17	0.529	0.610	36	0.384	0.454
18	0.514	0.594	37	0.381	0.450
19	0.501	0.580	38	0.377	0.446
20	0.489	0.567	39	0.374	0.442
21	0.478	0.555	40	0.371	0.438

**例 16—12** 用外标法定量，标准试样共进样 10 次，依次得到峰高（mm）如下：  
142, 146.5, 146.4, 146.3, 147.7, 135.0, 162.0, 140.0, 143.5, 146.3 在取平均峰高之前，  
检验一下哪些测定值要舍弃？

题解：①首先按由小到大排列：135.0；140.0；142；143.5；146.3；146.3；146.4；146.5；  
147.7；162.0 受检验的是两个端值。

②根据表 16—3 公式，计算

$$Q_{11} = \frac{Z(2) - Z(1)}{Z(H-1) - Z(1)} = \frac{140-135}{147.7-135} = 0.394$$

$$Q_{11} = \frac{Z(H) - Z(H-1)}{Z(H) - Z(2)} = \frac{162-147.07}{162-140} = 0.650$$

③从表 16—4 查出检验显著概率为 5% 和 1% 统计量的临界值为： $Q_{0.05(10)} = 0.530$

$Q_{0.01(10)} = 0.535$

④判定： $\because Q_{11} < Q_{0.05(10)} = 0.530 \therefore 135.0$  值正常接受。

而  $Q_H > Q_{0.01(10)} = 0.535$  因此 162.0 值为界外值，舍弃不要。

如果计算的统计量  $Q$  介于 0.530~0.635 之间，则为可疑值，但本组数据不在可疑值。

⑤舍去 162.0<sup>\*\*</sup>测定值后，还需检验 147.7 这一新的端值，就像重新提供 1 组测定值样，还需要重新算起，只是此时  $H=9$  即：

$$Q_{11} = \frac{Z(2) - Z(1)}{Z(H-1) - Z(1)} = \frac{140.0-135.0}{146.5-135} = 0.435$$

$$Q_{11} = \frac{Z(H) - Z(H-1)}{Z(H) - Z(2)} = \frac{147.07-146.5}{147.07-140.0} = 0.156$$

而查得  $Q_{0.05(9)} = 0.564$        $Q_{0.01(9)} = 0.672$

$$\because 0.435 < Q_{0.05(9)} = 0.564 \quad 0.156 < Q_{0.05(9)} = 0.564$$

故检验结果 135.0 及 147.7 均为正常保留值，应按 9 个计算平均值。

从上例不难看出，狄克逊检验准则拒绝接受的只是偏差很大的测定值，它把非异常值误判为异常值的概率是很小的，而把异常值误判为非异常值的可能性则大些。因而用狄克逊检验的数据，精密度不可能有偏高的假象，是一个比较好的检验方法。同时也使我们认识到，实验数据不能随意取舍。比如有人做了三次重复测定，往往有两个测定值比较接近，另一个数据有较大偏差。有的人则喜欢从三个测定值中挑选两个“好”的数据进行计算，另一个数据则丢弃不管。实际上，根据统计原理从三个数据中挑选两个是不合理的，不科学的，要纠正这种盲目行为。

## (2) 格布鲁斯 (Grubbs) 检验法

本法用于一组测量值或多组测量值的平均值的一致性检验和排除异常值，应用格布鲁斯检验时，按下述三种不同情况进行处理。

a、在只有一个可疑值的情况下：

①将几个测定值有小到大排成  $X_1, X_2, X_3, \dots, X_n$ ，设  $X_d$  被检验的可疑值（包括最大或最小值）；

②计算统计量  $G$

$$G = \frac{|X_d - \bar{X}|}{S} \quad 16-20$$

式中  $\bar{X} = \frac{\sum X}{n}$ ， $S$  为标准差； $X_d$  为被检验的最大或最小可疑值。

③查格布鲁斯检验临界值表 16-5，查出相应显著性水平  $\alpha$  和测定次数  $n$  时的临界值  $G_{\alpha, n}$

④判断：

当  $G \leq G_{0.05}$  则可疑值为极端值应保留

当  $G > G_{0.01}$  则可疑值为异常值应舍弃

若  $G_{0.05} < G < G_{0.01}$ ，则该值属技术原因产生的可以舍去，否则保留。

**例 16-13** 有 10 个实验室，对同一试样进行测定，每个室 5 次测定的平均值分别是 4.41, 4.49, 4.50, 4.51, 4.64, 4.75, 4.81, 4.95, 5.01, 5.39。检验最大均值是否为异常值？

题解：均值的平均值  $\bar{X} = \frac{\sum X_i}{n} = 4.75$

均值的标准差  $S_{\bar{x}} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} = 0.30$

$$\therefore G = \frac{|X_d - \bar{X}|}{S} = 2.13$$

查表 16-5；当  $n=10$  和显著性水平  $\alpha=0.05$  时， $G_{0.05, 10}=2.18$ ， $G < G_{0.05, 10}=2.18$ ；表明最大均值 5.39 为正常值。

b、如果可疑值有两个或两个以上，而且可疑值在同一侧，在检验时可以人为地暂时舍去两个可疑值的偏差更大的一个，用  $n-1$  个测定值计算平均值和标准差  $S$ ，检验偏差较小

的一个可疑值，若为异常值则先前舍去的必然为异常值，若检验值为不异常值，这时再由全部  $n$  个测定值计算平均值和标准差，去检验舍去的那个可疑值，根据检验结果确定是否为异常值，再决定取舍。

c、如果可疑值为两个或两个以上，并且分布在平均值两侧，检验方法同 b。

表 16—5 格布鲁斯检验临界值 G 表

n	显著性水平		n	显著性水平	
	0.05	0.01		0.05	0.01
3	1.15	1.15	17	2.47	2.78
4	1.46	1.49	18	2.50	2.82
5	1.67	1.75	19	2.53	2.85
6	1.82	1.94	20	2.56	2.88
7	1.94	2.10	21	2.58	2.91
8	2.03	2.22	22	2.60	2.94
9	2.11	2.32	23	2.62	2.96
10	2.18	2.41	24	2.64	2.99
11	2.24	2.48	25	2.66	3.01
12	2.29	2.55	30	2.74	3.10
13	2.33	2.61	35	2.71	3.18
14	2.37	2.66	40	2.87	3.24
15	2.41	2.70	50	2.96	3.34
16	2.44	2.74			

例 16—14 某实验室对同一试样进行 10 次测定的结果为：73.5，69.5，69.0，69.5，67.0，67.0，63.0，69.5，70.0，70.5。试问可疑值 63.5 与 73.5 是否为异常值。

题解：这是两个可疑值分布在平均值两侧的情况，测定平均值为 68.9，两个值偏差分别是 -5.4 和 +4.6，因此是暂时舍去可疑值 +63.5，用其余 9 个测定值去计算平均值  $\bar{X}$  和标准差  $S$ ，检验可疑值 73.5。

$$\text{这时：}\bar{X}=69.5, \quad S=1.9。G = \frac{|X_d - \bar{X}|}{S} = 2.11$$

查表 16—5；当  $n=9$  和显著性水平  $\alpha=0.05$  时， $G_{0.05, 9}=2.11$ ， $G < G_{0.05, 10}=2.11$ ；表明可疑值 73.5 不能作为异常值舍弃，应该保留。

$$\text{再用 10 个测定值计算：}\bar{X}=68.9, \quad S=2.6。G = \frac{|X_d - \bar{X}|}{S} = 2.1$$

查表 16—5；当  $n=10$  和显著性水平  $\alpha=0.05$  时， $G_{0.05, 10}=2.18$ ， $G < G_{0.05, 10}=2.18$ ；表明可疑值 63.5 也不能作为异常值舍弃。

### 三. 测定结果的校正

在食品分析中常常因为系统误差，使测定结果高于或低于检测对象的实际含量，即回收率不是 100%，所以需要在样品测定时，用加入回收法测定回收率，再利用回收率按下式以样品的测定结果的校正。

$$X\% = X_0\% \div P\%$$

16—21

式中：X%—样品中被测成分的百分含量；

$X_0\%$ —样品中被测成分测得的百分含量；

$P\%$ —回收率。

思考题：

1. 当有一组测量值，其总体标准偏差  $\sigma$  为未知，要判别得到这组数据的分析方法是否可靠，应该性质下列方法中的哪一种？  
(1) 4d 法    (2) 格布鲁斯法    (3) F 检验法    (4) t 检验法
2. 回收率在在一定程度上决定了方法的可靠性，一般可靠的回收率范围是多少？
3. 当测定次数很多时，下列关于标准偏差  $\sigma$  与平均偏差  $\delta$  之间关系是什么？
4. 两组分析人员对同一含  $\text{SO}_4^{2-}$  的试样用重量法进行分析，得到两组分析数据，要判断两人分析的精密性有无显著性差异，应该用下列哪一种方法？  
(1) Q 检验法    (2) F 检验法    (3) u 检验法    (4) t 检验法
5. 在上题中，若要判断两分析人员的分析结果之间是否存在系统误差，则应该选用下列哪种方法？  
(1) u 检验法    (2) F 检验法    (3) F 检验法加 t 检验法    (4) t 检验法

(章银良)