数理统计学在评价食品分析方法中的应用

安徽省食品科技服务中心　赵齐川

食品营养价值的高低可通过测定食品中某些营养成分的含量来确定。分析测定食品营养成分含量的方法甚多，不同的营养成分有其不同的分析方法，同一种营养成分可采用多种方法进行定量。以抗坏血酸为例，它在食品中的含量可采用微生物测定法；2,4—二硝基苯肼比色法；钼兰比色法；22′—双吡啶比色法；荧光比色法；高效液相色谱法，碘量法和2,6—二氯酚靛酚滴定法等方法进行定量。

随着科学技术的发展，检测仪器的更新，分析手段的完善，可以预见，今后将会建立起更多更好的分析方法来。

要建立一种新的方法是一项科学性很强的工作，需要做很多的细致工作，其中必不可少的工作之一，是应用数理统计学的原理和方法，对新法测定样品的结果进行计算，然后根据计算结果对新法作出准确的评价。

(一)准确度的应用

准确度是指某种测定方法测出样品某种成分的数值与其真实值的符合程度，可用公式 $\delta = m - M$ 表示(式中 δ 为准确度，m 为分析值，M 为真实值)。一种好的测定方法，用于测定样品时，测得的数值应与样品中实际所含该成分的真实一致或接近。

判断一种方法的准确程度是把该法测得某种成分的结果与其真实值进行比较。通常是采用新法测定已知浓度的标准液，用其测定结果与标准液真实值进行比较。然而食品中某种成分的真值是无法知道的，在这种情况下，一般采取测定其回收率来确定准确度。所谓回收率就是取一定份数的待测样品，并就每份样品再分成两部分，其中一部分另外加入已知量的被测物质，然后用所研究的测定方法同时对两者进行测定，将测得的结果按回收率公式进行计算。

$$\text{回收率} = \frac{\text{已知量被测物质的测定结果} - \text{样品的测定结果}}{\text{被测物质的量}}$$

回收率高说明该法准确度高。现以《食品科学》87年6期58页表1为例，作者采用钼兰比色法测定维生素 C，其回收率 94～102%。结果说明此法用于测定维生素 C 含量是可行的。

当然在实际工作中没有绝对准确的方法，只是要求所使用的方法要有一定的准确度，决不能把其准确度太差的方法用于食品定量分析，以避免对实验结果作出不正确的评价。

二　精密度的应用

表示几次测定结果与测定平均值的偏差程度的量称为精密度。精密度的大小通常用标准差或差异系数来衡量。一种高精密度的测定方法，其重复测定某一样品时，几次测定结果的差异是很小的。

1. 同标准差表示精密度

标准差计算公式如下

$$S = \sqrt{\frac{\Sigma(X - X_{均})^2}{n-1}}$$

式中 S 的标准差；X 为各测定值；n 为重复测定次数；Σ 读作 Sigma，表示总和；$X_{均}$ 为平均数，计算公式为 $X_{均} = \frac{\Sigma X}{n}$。

标准差是度量一群性质相同数值变动范围大小的指标。采用某种方法重复测定某一样品，计算出的 S 值越小，表明这种方法的精确度越高。

例如以一份已知浓度的标准抗坏血酸采用某种方法进行5次重复测定，其结分别为(mg%)：5.3、5.7、5.4、5.2、5.6，试计算标准差。

根据题意和上述公式可求出 $\Sigma X = 27.2$，

$X_{均}=5.44$，$\Sigma X^2=148.14$，$n=5$，$S=0.207364$ mg%。

2. 用差异系数表示相对精密度

差异系数是由标准差除以平均数算得的百分比，差异系数用C、V代表，其计算公式为 $C、V=\frac{S}{X_{均}}$

在什么情况下使用差异系数表示相对精密度呢？为了说明这个问题，现借用《食品科学》87年6期58页表2的数据。作者采用染料滴定法测得刺梨可乐（329号样品）中维生素C平均含量为10.13，标准差为2.23mg/100ml；钼兰比色法测得浓刺梨汁（245号样品）中维生素C平均值为295.8，标准差为30.3mg/100ml。单就标准差而言滴定法远远小于钼兰比色法，但由于两个不同样品中维生素C含量差异很大，很难根据标准差判断哪种方法好。在这种情况下我们采用标准差除以各自的均数算得百分比来表示相对精密度。在我们这个例子中用染料滴定法测定浓刺梨汁维生素C的差异系数为22.01%；钼兰比色法测定刺梨可乐维生素C的差异系数为10.24%。通过这样计算，可以看出钼兰比色法要比染料滴定法精确度高。

三 平均差数的显著性检验

对于一种新建立的方法需要知道它在检测中是否适用，如何达到这一目的，其中重要的一条就是用新法和已公认的原法对同一批样品每份各测定一次，然后求得两种方法测得结果的差数的均数，并通过平均差数的显著性检验——t检验后做出正确判断，若两种方法差异不显著，认为新法可行，若存在着显著性差异（即 $P\leqslant0.05$）说明新法尚有一定缺陷。

现举例说明平均差数的显著性检验的计算方法。取10个样品，对每一样品用新法和原法各测一次抗坏血酸含量其结果如表1。

表1

新法	381	627	485	546	516	556	595	436	569	595
原法	376	620	494	563	480	598	543	383	487	512
d	5	7	−9	−17	36	52	52	53	82	83

根据测定结果先计算两种方法测定各样品差数d（计正负号）、d^2和Σd、Σd^2以及$d_{均}$（$d_{均}=\frac{\Sigma d}{n}$）然后用

$$Sd=\sqrt{\frac{\Sigma(d-d_{均})^2}{n-1}}=\sqrt{\frac{\Sigma d^2-\frac{(\Sigma d)^2}{n}}{n-1}}$$

公式求出$Sd_{均}$再用$t=d_{均}/Sd_{均}$求出t值，并通过t值表查自由度n'、$t_{0.05}(n')$相对应的t值，最后对计算出的t值与t值表查出的值进行比较，如算得的t值大于t值表相对应的值则$P<0.05$，说明两法差异显著，也就是说新法不理想；如计算的t值小于t值表相对应值则$P>0.05$，表明两种方法无差异，故新法可靠性较好。

本例$n=10$；$n'=n-1=9$；$\Sigma d=250$；$\Sigma d^2=22630$；$d_{均}=25$；$Sd_{均}=13.49$；$t=1.853$；$t_{0.05(9)}=2.262$，从计算结果看$t<t_{0.05(9)}$，说明新法测定抗坏血酸含量是可靠的。

四 相关系数

判定新建立的方法在实际检测中是否适用，也可通过对统计指标——相关系数的计算加以证明。新的测定方法如果适用的话，用该法和原有方法同时分别测定每份样品，其结果会存在着一一对应的关系。如果两种方法测定结果用点代表时，则在坐标系上得到一群（X，Y）点（每点表示一份样品的两种方法测定结果），此等点代表的图形（如图1）通常被称做散布图。图上的各点如在一条直线上时则相关系数为1（图1a）；如两种方法测定结果完全无关，则相关系数为0（图1b）。在研究食品同一营养成分的两种测定方法时，相关系数为0的情况一般是不会有的；相关系数为1的情况也很难出现，多数情况是相关系数接近1，相关系数越接近1，则相关程度越好。

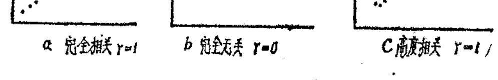

图1 新法、原法测定结果相关程度散布图

现以A法的横轴X（一般皆以标准方法作为横轴），B法为纵轴Y，计算两种方法所测得样品抗坏血酸含量的相关系数。

1.列计算表如表2。

表2

X（A法）	Y（B法）	X^2	Y^2	XY	X+Y	$(X+Y)^2$
19.5	19.3	380.25	372.49	376.35	38.8	1505.44
295.7	295.8	87438.49	87497.64	87468.06	591.5	349872.25
15.4	15.1	237.16	228.01	232.54	30.5	930.25
10.13	13.6	102.617	184.96	137.768	23.23	563.1129
0.68	2.20	0.4624	4.84	1.496	2.88	8.2944
1.04	2.3	1.0816	5.29	2.392	3.34	11.1556
37.5	41.1	1406.25	1689.21	1541.25	78.6	6177.96
35.7	39.4	1274.49	1552.36	1406.58	75.1	5640.01
415.65	428.8	90840.799	91534.8	91166.436	843.95	364708.46

表中X、Y为两种方法测定结果，X^2，Y^2为各测定值平方，XY为X与Y的乘积，X+Y为X与Y之和，$(X+Y)^2$为(X+Y)的平方，各列相加得总和于表末行。

2.根据公式计算相关系数

$$相关系数=\frac{\Sigma XY-(\Sigma X)(\Sigma Y)/n}{\sqrt{[\Sigma X^2-(\Sigma X)^2/n][\Sigma Y^2-(\Sigma Y)^2/n]}}$$

$$=\frac{91166.436-415.65\times428.8/8}{\sqrt{[90840.799-(415.65)^2/8][91534.8-(428.8)^2/8]}}$$

$$=\frac{68887.596}{68897.274}=0.999$$

从计算结果看，相关系数与1非常接近，说明新建立的方法与原有标准方法测定结果关系十分密切，足以证明新法是适用的。

五 回归分析

为了了解新法与原法之间的确切关系，当各测定结果在相关图上呈直线关系时，可以算出一条最能代表这些点子分布趋势的直线方程式来表示两种测定结果的关系。这一方程式称为回归方程。

回归方程的计算是利用相关分析表的结果，用下式求出回归方程 Y=a+bX 的斜率b及截距a。

$$b=\frac{\Sigma XY-(\Sigma X)(\Sigma Y)/n}{\Sigma X^2-(\Sigma X)^2/n}$$

$$a=Y_{均}-bX_{均}$$

式中$Y_{均}$、$X_{均}$为Y及X值的均数。

本例用求相关系数时算得的数据代入得：

$$b=\frac{68887.596}{69245.184}=0.9948$$

$$a=53.6-51.95625\times0.9948=1.912$$

由此得回归方程 $\hat{Y}=1.912+0.9948X$。式中 $\hat{Y}$ 表示由X推算得的Y值，以与实际测定结果Y值相区别。

回归方程$\hat{Y}=a+bX$，a为截距，就是当X为零时回归线在纵轴上的高度，如果两种测定结果完全相同，回归线应通过原点，即a=0。b是回归线的斜率，由于X与Y的单位相同，所以b就是图中Q角的正切。当b=1时表示新法与原法增加量成正比。

从回归分析本例结果来看，新法与原法关系很好，可以说新法适用于测定样品中成分的含量。

t 值 表　　表3

自由度n'	t0.05(n)'	t0.01(n)'	自由度n'	t0.05(n)'	t0.01(n)'
1	12.706	63.657	20	2.086	2.845
2	4.303	9.925	30	2.042	2.750
3	3.182	5.841	40	2.021	2.704
4	2.776	4.604	50	2.008	2.678
5	2.571	4.032	100	1.984	2.626
6	2.447	3.707	500	1.965	2.586
7	2.365	3.499	1000	1.962	2.581
8	2.306	3.355	∞	1.960	2.576
9	2.262	3.250			
10	2.228	3.169			